Residential Segregation and Interracial Friendship in Schools[1]

Ted Mouw and Barbara Entwisle
University of North Carolina, Chapel Hill

This article uses social network and spatial data from the National Longitudinal Study of Adolescent Health (Add Health) to examine the effect of racial residential segregation on school friendship segregation in the United States. The use of hierarchical models allows the authors to simultaneously estimate the effects of race, within-school residential segregation, and school diversity on friendship choice using the Add Health data. The authors use these results to predict the decline in friendship segregation that would occur if across- and within-school residential segregation were eliminated in U.S. metropolitan areas. The results suggest that about a third of the level of racial friendship segregation in schools is attributable to residential segregation. Most of this effect is the result of residential segregation across schools rather than within them.

INTRODUCTION

Three decades after the end of legalized racial segregation in the 1960s, the social worlds of black and white Americans are still largely separate. Residential segregation between blacks and whites has declined over the past 30 years (Glaeser and Vigdor 2001; Cutler et al. 1999), but high levels of segregation are still a defining feature of most urban areas in America (Massey and Denton 1993; Farley and Frey 1994; White 1987). The res-

[1] The authors are grateful for the support provided by the Carolina Population Center of the University of North Carolina, Chapel Hill. They would also like to thank Michael White, Lincoln Quillian, Richard Alba, Yu Xie, and Katherine Faust for their comments and advice on an earlier draft of the article. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Direct correspondence to Ted Mouw, Department of Sociology, University of North Carolina, Chapel Hill, CB#3210, 155 Hamilton Hall, Chapel Hill, North Carolina 27599-3210. E-mail: tedmouw@email.unc.edu

0002-9602/2006/11202-0002$10.00

idential segregation of Hispanics and Asians from whites, although less extreme than that of blacks, appears to be increasing (Charles 2000).

Patterns of social or friendship segregation seem to parallel those for residential segregation. Although there is evidence of increasing social contact between whites and blacks (Sigelman et al. 1996), genuine interracial friendship between blacks and whites still seems to be the exception rather than the rule (Shipler 1997; Johnson and Marini 2000; Marsden 1987; Sigelman et al. 1996). Recent data indicate that while a substantial number of whites and blacks claim to have interracial friends (Smith 1999; Sigelman and Welch 1993; Sigelman et al. 1996), when asked to list the names of their close friends, only 6% of whites and 15.2% of blacks actually listed a friend of the other race (Smith 1999). We know less about how such patterns extend to other race and ethnic groups in America, although it does appear that Hispanic and Asian adolescents are also very likely to choose as friends other members of their same racial or ethnic group (Quillian and Campbell 2003).

The available evidence documents high levels of both residential segregation and friendship segregation, but is there a link between them? If, because of residential segregation, neighbors overrepresent a particular racial or ethnic group, then friends may likewise overrepresent this group. The literature on residential segregation shows the most segregation between blacks and whites, with Hispanics and Asians in an intermediate position (Charles 2003). The literature on friendship choices shows a parallel pattern (Quillian and Campbell 2003). The correspondence between the two is consistent with a link, but it falls far short of providing convincing evidence for a connection between residential and social segregation. For this, we need information about friendship choice linked to residential location that would allow researchers to investigate this relationship at the appropriate scale.

In this article, we use unique data from the National Longitudinal Study of Adolescent Health (Add Health) to analyze the effect of residential segregation on within-school friendship patterns among diverse race and ethnic groups.[2] The "in-school" portion of the Add Health survey collected data on friendships. The "in-home" portion of the Add Health survey collected data on the spatial location of respondents' homes, from which we can calculate the geographic distance between pairs of potential friends. Combining these two sources of data, we can see whether in-

[2] For simplicity, we refer to race and ethnic categories collectively as "racial" categories. We recognize that the term "Hispanic" more closely corresponds to what sociologists have traditionally meant by ethnicity than race. However, these traditional distinctions have ceased to have meaning as both race and ethnicity have become understood as broad social categories related to "origins" (Hirschman et al. 2000).

creasing distance between pairs of potential friends decreases the likelihood of friendship between them and test the degree to which residential segregation by race affects friendship segregation.

The use of residential proximity measures makes it possible to get beyond some of the limitations of census geography. For instance, the literature on segregation generally shows increasing levels of segregation with decreases in the size of the unit of measurement (Charles 2003). The use of residential proximity measures provides an opportunity to examine patterns within as well as between census units such as tracts, the usual unit for studies of residential segregation. It also allows us to avoid arbitrary edge effects, for example, when persons on opposite sides of a tract boundary are considered to come from different neighborhoods whereas persons at opposite ends of the same tract are treated as coming from the same neighborhood. Indeed, as we will show, residential proximity effects are quite local.

We estimate a hierarchical model of friendship choice that takes into account school racial composition as well as race, residential proximity, and other characteristics on the choices of potential friends within the school. The effects of residential proximity, and changes in the effects of race when residential proximity is added to the model, capture the impact of within-school residential segregation. To consider the impact of between-school residential segregation, we use our results to simulate the decline in levels of friendship segregation that would occur in each U.S. metropolitan area in the absence of residential segregation, using data from the National Center for Educational Statistics (NCES). The racial segregation of public schools in the United States is high (Orfield and Lee 2004, p. 17), and recent evidence suggests that black and Latino students are becoming more segregated from white students in most areas of the country (Frankenberg and Lee 2002).

Although the focus of this article is primarily substantive, we also address the implications of the problem of statistical nonindependence for studies of interracial friendship. This problem arises if individuals' friendship choices are affected by the choices of their friends—if, for example, two individuals are more likely to become friends if they have a mutual friend in common. The consequence is that friendship data are endogenous, and estimates of the effect of race may be biased. We use a Monte Carlo simulation to show that conventional statistical methods may substantially overestimate the degree of same-race preference if friendship is also affected by factors such as residential location and social class that are correlated with race, even if these variables themselves are also included in the models. Although at the moment there is no perfect remedy to the problems posed by the endogeneity of network structure, we use both a conventional logit model as well as an exponential random graph

(p*) model that incorporates data on network structure, such as the number of mutual friends between individuals, to show that our results are robust to different methodological approaches.

DETERMINANTS OF SOCIAL SEGREGATION IN SCHOOLS: OPPORTUNITY AND PREFERENCE

The level of social segregation in any given context—for example, in a neighborhood, in school, or at work—may be thought of as the result of two basic factors: propinquity, the opportunity for interracial contact; and homophily, the preference of individuals to associate with others who are similar to them (Blau 1977). Opportunity for contact reflects the basic demographic or numerical distribution of groups in a particular context, and the structure and organization of that context. In the absence of any preference for same-group contact, social segregation would be determined by the random chance of cross-group interaction. In a macrosociological theory based upon this principle, Blau (1977) observes that while individuals may prefer to associate with members of their own group, the relative size of their group influences their ability to satisfy those preferences. In addition, within a given context, there may be structures that encourage or discourage contact between certain individuals. In neighborhoods, these might include general spatial layout, the organization of tertiary streets, whether or not houses face each other, and the existence of focal points for interaction such as bus stops, coffee shops, and the like (Grannis 1998). In schools, these might include tracking and the way that students are assigned to classes, scheduling, and extracurricular activities on-site including clubs and sports (Hallinan and Williams 1989; Moody 2001). The existing evidence suggests that propinquity—i.e., the opportunity to interact—is an important factor in interracial friendship (Sigelman et al. 1996; Johnson and Marini 2000). Propinquity itself does not guarantee social integration, of course. Preferences for same-group homophily may result in social segregation even in racially mixed environments. In an extreme case, Falah (1996) shows that there is little social interaction between Arabs and Jews in mixed Israeli cities even when there are high levels of residential integration. Overall, the social-psychological evidence suggests that individuals tend to choose friends similar to themselves on a wide variety of characteristics such as social class, gender, age, race, and attitudes (Kandel 1978; Shrum et al. 1988; McPherson et al. 2001).

There is a large empirical literature on the causes and consequences of racial friendship segregation in general (Jackman and Crane 1986; Ellison and Powers 1994; Mayhew et al. 1995; Way and Chen 2000; Sigelman et al. 1996; Verbrugge 1983) and in schools (Quillian and Campbell 2003;

Kubitschek and Hallinan 1998; Hallinan and Williams 1987; Hallinan and Teixeira 1987; Tuma and Hallinan 1979; Fisher and Hartmann 1995; Moody 2001; Lubbers 2003; Dutton, Singer, and Devlin 1998; Damico and Sparks 1986; Joyner and Kao 2001; Schofield 1982; Shrum, Cheek, and Hunter 1988). One advantage of studying social segregation in schools is that the researcher can gather information on the distribution of potential friends in the school. This allows the researcher to differentiate between propinquity and homophily effects by constructing dyad data, where the cases comprise all possible pairs of students in the data who attend the same school. The structure of dyad data controls for the effect of propinquity at the school level because each case represents a potential friend, and homophily can be tested by seeing whether demographically similar dyads are more likely to be friend dyads.[3] Hallinan and Williams (1989) use dyad data on pairs of potential friends in the same class to study interracial friendship using High School and Beyond data. Overall, they find a significantly lower probability of friendship for black-white dyads even after controlling for other factors such as tracking, shared activities, and class rank.

The advent of the National Longitudinal Study of Adolescent Health has greatly facilitated research on relationships among adolescents in general (Giordano 2003), and interracial friendship patterns in particular. Joyner and Kao (2001), Moody (2001), and Quillian and Campbell (2003) each take a different approach to studying racial friendship homophily with the Add Health data set. Using data on the race of the respondent's best friend, Joyner and Kao (2001) find that the tendency toward same-race friendship weakens as the composition of the school becomes more diverse. However, because Joyner and Kao use individuals as cases rather than dyads of potential friends, their results combine the effects of propinquity and homophily. A great advantage of the Add Health data is the availability of data on potential as well as actual friends.

Moody (2001) estimates the effect of school diversity on racial homophily using a two-stage strategy. In the first stage of his analysis, Moody uses dyad data on potential friends to estimate the odds of same-race friendship in separate models for each of the 134 schools in the Add Health data, using p^* models that attempt to control for the endogeneity of network structure. In the second stage, he estimates the effect of school racial diversity on the estimated odds of same-race friendship from the first stage. Because he uses dyad data on potential friends in the first stage, Moody controls for the propinquity effect of school racial composition.

[3] However, see our depiction of "network" propinquity below for a discussion of how opportunity for friendship may be constrained even among students who go to the same school.

Moody finds that higher racial diversity decreases the probability of cross-race friendship for any given pair of potential friends of different races. This suggests the existence of "school climate" effects as a result of heightened racial awareness or tension in racially diverse schools. One drawback of Moody's aggregate approach, however, is that it collapses the multiracial Add Health friendship data into a single indicator of same- or cross-race friendship.

Quillian and Campbell (2003) use Hallinan and Williams's (1989) dyad-based analysis of potential friends and extend their analysis by studying multiracial friendship segregation among black, white, Hispanic, and Asian students using the Add Health data. Of key interest in this study is whether Hispanic and Asian friendship preferences weaken across immigrant generations, as predicted by traditional assimilation theory, or remain segmented. Quillian and Campbell (2003) find that similarity in racial background is a powerful factor influencing adolescent friendship choices for Hispanic and Asian adolescents, as it is for white and black ones. They find little evidence for a weakening of racial homophily across generations. However, because they do not take the nonindependence of the dyads into account it is possible that their results overestimate the degree of racial homophily in the data. This point will be further discussed in the next section of the article.

In this article, we go beyond earlier studies based on Add Health data in considering a propinquity effect whose source lies outside of the school setting, in residential neighborhoods. Social processes occurring within schools reflect influences originating outside as well as inside them. The existing evidence suggests that spatial proximity is an important factor in friendship formation. In a study of a housing community, Festinger, Schachter, and Back (1950) show that physical proximity is an important determinant of passive social contact among neighbors. Similarly, neighborhood proximity affects friendship formation and the frequency of contact (Fischer et al. 1977; Huckfeldt 1983), and the median distance from friends with whom we discuss important matters is small (less than 10 miles; Latane et al. 1995). In contrast to these studies, however, the study of the effect of space on school friendship is different because joint school attendance provides an institutional context for interaction in addition to neighborhood proximity. In a detailed study of 10 high schools in Indiana, Patchen (1982) reports evidence of both propinquity and homophily effects on interracial friendship. The percentage black in students' classes and the level of black-white neighborhood contact affects the probability of cross-race friendship, as do positive family attitudes, educational aspirations, and the school's racial climate. Similarly, DuBois and Hirsch (1990) study the relationship between school friendship and neighborhood friendship patterns among blacks and whites at an integrated junior high school and find that students who lived in neighborhoods with more other-race children are more likely to report having at least one close friend of the other race. We seek to build upon these studies by testing the effect of residential proximity on school friendship segregation using network data from a larger sample of schools and precise information on residential location. Our methodological approach is similar to Lubbers (2003), who estimates multilevel p^* logit models of friendship segregation of immigrants using a two-step method with dyad data on within-class friendships for school children in 57 school classes in the Netherlands. Rather than a two-step approach, however, we use hierarchical models to simultaneously estimate the effects of individual- and school-level variables.

FRIENDSHIP SEGREGATION: RACE, SPACE, AND THE PROBLEM OF NETWORK ENDOGENEITY

How can we differentiate the effect of preferences for same-race friendship from other factors that may contribute to friendship segregation? The usual approach identifies, measures, and takes into account factors related to propinquity and homophily on characteristics other than race, interpreting whatever remains as a racial homophily effect. This indirect or residual approach to examining racial homophily is vulnerable to many kinds of model misspecification, including incomplete measurement of propinquity effects. In this section we discuss the interconnection between race, residential segregation, and network structure and show why the problem of the nonindependence of network data has important substantive implications for the study of racial homophily in general and in our particular study. Specifically, we show that if not properly handled, nonindependence can result in overstatement of the homophily effect.

An Example

We begin with an example. "Diversity High School" is the pseudonym for a school in the Western part of the United States that was sampled as part of the Add Health data. Diversity High is interesting because it is an almost all-minority school composed of roughly equal proportions of blacks, Asians, and Hispanics. According to the self-reports of the students in the school, approximately 5% are white, 20% are black, 35% are Asian, and 40% are Hispanic.[4] However, despite this high degree of diversity in the composition of the school, friendship patterns in Diversity

[4] For clarity of presentation, we exclude a small number of Native American students and students with a race as missing or other.

High School are very segregated. As shown in table 1, 75.5% of blacks' friends are black, 86.4% of Asians' friends are Asian, and 82.4% of Hispanics' friends are Hispanic. Only whites, who are a small minority of the school, show any evidence of racial integration in their friendship patterns. Even in the absence of a white majority exhibiting a tendency toward same-race friendship (as is evident in other Add Health schools), blacks, Asians, and Hispanics display high levels of same-race homophily.

Although table 1 shows high levels of racial friendship segregation, it might not all arise from a preference toward same-race friendship. Some of the friendship segregation may be the result of underlying levels of residential segregation and socioeconomic inequality. Indeed, Diversity High School has a distinctive pattern of residential segregation. Figure 1 shows the location of students' houses, coded by race and ethnic group.[5] Visually, it is evident that black students are, for the most part, segregated from their Asian and Hispanic classmates. Neighborhood 1, on the left, is a predominantly black neighborhood, while neighborhood 2, on the right, is mostly Asian and Hispanic, with a small number of black households. Descriptively, at least, residential location can be shown to affect friendship choice: of Asians and Hispanics living in the predominantly black neighborhood 1, 18.8% of Asians' friends, and 13.9% of Hispanics' friends are black, compared to 2.3% and 3.2%, respectively, of those not living in neighborhood 1. For blacks living in the largely Asian and Hispanic neighborhood 2, 35% of their friends are nonblack, versus only 18% of those not living in neighborhood 2. In addition to residential segregation, in Diversity High School there is a clear class difference between the Asian and black students, on the one hand, and the Hispanic students, on the other. The average parental income and education of the black and Asian students is considerably higher than that of the Hispanic students. If young people prefer to make friends with others who share their class background, these patterns alone could explain an apparent preference for same-race friends.

The above discussion of figure 1 and table 1 helps to illustrate the uniqueness of the Add Health data. Information on friendship choices combined with data on residential location and class differences can help explain the role of space and class on friendship segregation. In Diversity

[5] Map 1 shows the residential location of Diversity High School students living in two geographic areas. Although the spatial relationships among these residences are preserved, the areas have been cropped and moved to preserve confidentiality. Although we refer to these areas as "neighborhoods" in the text, they do not correspond to census tracts. As a consequence, it is not possible to recover group totals from the map or to link information from the map, the text, and table 1 to identify the school. Map 1 has been reviewed and approved by the Add Health team (Kathleen Mullan Harris, principal investigator).

TABLE 1
RACIAL FRIENDSHIP SEGREGATION IN "DIVERSITY HIGH SCHOOL"

RACE OF RESPONDENT	RACE OF FRIEND (Row Percentage)				
	White	Black	Asian	Hispanic	Total
White	20.9	5.5	25.3	48.4	100
Black	.8	75.5	7.6	16.1	100
Asian	1.0	2.7	86.4	9.9	100
Hispanic	3.1	3.6	11.0	82.4	100

High, high levels of residential segregation and class differences may contribute to the high level of friendship segregation evident in table 1. This poses a basic question: How much friendship segregation would there be if we controlled for residential segregation and race differences in social class? Or, putting it another way, how much do patterns of residential segregation limit social integration in the schools? Before we turn to a more systematic analysis of this question, however, we must consider how the endogeneity of friendship networks—i.e., how our choice of friends is affected by our friends' choice of friends—may affect our results. It turns out that it is more difficult to "control" for other factors such as social class and residential segregation in social network data than in conventional data sets. This is a problem that we need to solve for our own purposes, but which also has implications for other analyses done with this and similar data sets.

Mutual Friends and the Endogeneity of Friendship Networks: A Simulation

Imagine a hypothetical white adolescent who moves with her family to a town that is integrated in terms of its overall composition of blacks and whites, but residentially segregated. The school she is districted to attend is predominantly white. In addition, assume that the students who already attend the school have friends who are mostly white. Even if our hypothetical adolescent has no preference toward same-race friends, the initial friends she meets at school will be mostly white. Likewise, because of residential segregation the people she meets first in her neighborhood—walking down the sidewalk, hanging out in the park, etc.—are likely to be white. As time passes, the friends she meets via her initial friends—through countless formal introductions and informal social interactions—will also be mostly white. Clearly, our hypothetical adolescent is likely to end up with mostly white friends even if she has no underlying preference one way or the other.

Given such a situation, how can we differentiate between the observed

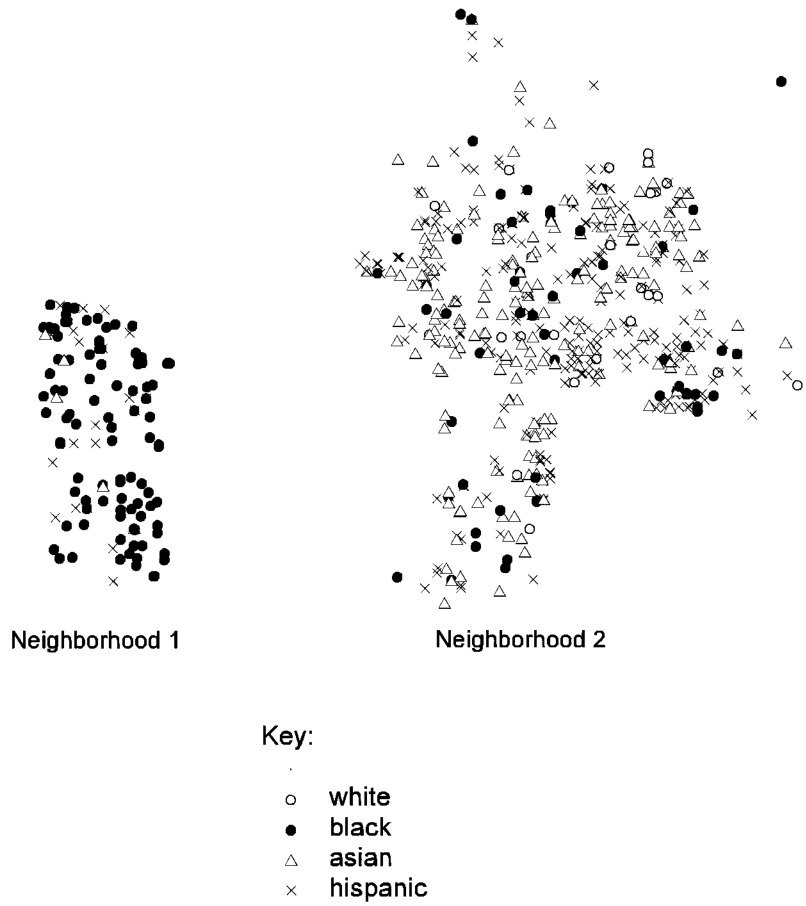

FIG. 1.—Residential segregation in “Diversity High School”

level of friendship segregation and the underlying “preference” for racial homophily? It is not just that there are other observable factors such as residential location and class differences that affect friendship choices, but that friendship choices are interdependent—the people we come into contact with and are introduced to are not a random sample of the town that we live in or school that we go to, but are constrained by the friendship choices of our friends, who in turn are constrained by the friendship choices of their friends, and so on.

The basic result of the interconnectedness of friendship choices is that friendship data are not statistically independent. Social network theorists have described the way in which this dependency may express itself through network structure and have attempted to develop methods to overcome the problem of nonindependence (Frank and Strauss 1986; Strauss and Ikeda 1990; Wasserman and Pattison 1996; Pattison and Wasserman 1999; Snijders 2002). We implement some of their recommendations. Before going into this methodological literature, however, it is useful to consider the substantive significance of network endogeneity on the analysis of friendship segregation.

The simplest way to think about the endogeneity of network structure is to consider the role of mutual friends. If you are more likely to become friends with someone if you have mutual friends in common (i.e., you may meet them through the mutual friends), then your choice of friends is affected not only by your individual preferences, but also by the friendship choices of your friends. Hence, an initial preference for racial homophily will be magnified in the resulting network, as your choice set of potential friends is not a random sample of the population but a reflection of the preferences of your friends. This can be thought of as a "network" propinquity effect because your opportunities for interracial friendship are constrained by the friendship choices of your friends.

If race were the *only* exogenous factor affecting friendship choice, then the observed level of racial friendship segregation would simply combine the direct effect of race via homophily preferences with the indirect effect of race from endogenous network effects (which would be the result of other people's preferences for racial homophily). However, in the presence of other factors that affect friendship choice, such as residential segregation and social class differences, to name two, the situation gets more complicated. If, for example, friendship choices are partly based upon residential location, and residential location is correlated with race, then the endogenous role of mutual friends transmits the effects of both racial homophily and residential segregation. In contrast to data where the cases are independent, this will upwardly bias the coefficient on race even when we include data on residential location in the same model.

In order to develop an intuitive understanding of the role of mutual friends on racial friendship segregation, we simulate the evolution of social networks over time through race, residential segregation, and mutual friends.[6] The point of the simulations is to work backward from known conditions to see how much bias would result in our estimates of racial homophily using conventional logit models that ignore the endogeneity of network structure. The following simulation focuses on residential segregation because that is the subject of our article, but the same logic would hold for the effect of social class on friendship segregation. The

[6] The full results of these simulations and all the Stata programming files are available on request.

basic result is that conventional logit models may substantially overestimate the effect of race when there is dependency in friendship choices of different individuals and when there are other factors correlated with race that also affect friendship.

In the simulation we use a hypothetical school composed of 36 students, equally divided between blacks and whites. We assume the students live along a 6×6 grid; each point on the grid is separated from the next by .25 km. The black students live in the southwest and northeast quadrants of the school catchment area, and the white students live in the northwest and southeast quadrants.

We simulate the friendship networks for this school under a variety of conditions in scenarios 1–4. To do so, we first form dyads for all combinations of potential friends in the data. Each of the 36 students has 35 potential friends, so there are $36 \times 35 = 1{,}260$ dyads in the data. For each of the four scenarios, the social network data was simulated 100 different times, resulting in 126,000 dyads. Exogenous variables are created indicating whether or not the two students are the same race and whether they live within .25 km of each other. Each simulation consists of 10 iterations. During each of the 10 iterations, an endogenous variable measuring the number of mutual friends between two individuals in the previous time iteration is calculated; increasing the number of mutual friends connecting two students increases the probability that they will be friends. Each student is constrained to have five friends at any time in each iteration; the individual's five "best" friends are selected based upon the predictive model and an error term. The probability of friendship between two students is

$$\ln\left(\frac{\text{prob}[Y = 1]}{\text{prob}[Y = 0]}\right) = C + B_1(\text{same race}) + B_2(\text{live within } \tfrac{1}{4} \text{ km}) + B_3(\text{mutual friends}). \qquad (1)$$

After data have been simulated for 10 time iterations, we analyze the resulting data using three models to see how close our statistical estimates come to the "true" coefficients of the predictive model.

Table 2 shows the results from these simulations. Model 1 shows the results when race is the only independent variable in a logit model predicting friendship. Model 2 shows the results when all the exogenous factors are included in the regression model (but mutual friends are not). Model 3 is a pseudo likelihood "p^*" model which includes all the relevant exogenous variables plus the variable measuring the number of mutual friends connecting the pair of individuals (see appendix A). To begin with, scenario 1 shows what would happen if the data were independent. In scenario 1, there is a preference for same-race friendship ($B_1 = .6$) and a

TABLE 2
RESULTS FROM SOCIAL NETWORK SIMULATIONS

	Scenario 1 (Independence)	Scenario 2	Scenario 3	Scenario 4
Actual model:				
Exogenous factors	.6 × same race, 1 × (live ≤ .25 km)	.6 × same race	.6 × same race, 1 × (live ≤ .25 km)	.6 × same race, 1 × (live ≤ .25 km)
Endogenous factors		.6 × mutual friends	.6 × mutual friends	.5 × have friend within .25 km
Simulation results:				
Model 1: Bivariate				
Same race	.714	.713	1.105	1.301
	(.017)	(.017)	(.018)	(.018)
Constant	−2.186	−2.184	−2.436	−2.576
	(.013)	(.014)	(.014)	(.015)
Model 2: All exogenous factors				
Same race	.571		1.00	1.175
	(.017)		(.108)	(.019)
Live ≤ .25 km apart	1.040		.740	.961
	(.022)		(.023)	(.022)

Constant	−2.242		−2.470	−2.629
	(.013)		(.015)	(.015)
Model 3: All factors (p^* model, see text)				
Same race	.571	.581	.607	.648
	(.071)	(.017)	(.019)	(.020)
Live ≤.25 km apart	1.040		.914	.965
	(.022)		(.025)	(.024)
Mutual friends	−.0001	.526	.538	−.009
	(.006)	(.005)	(.005)	(.006)
Have friend within .25 km				.470
				(.005)
Constant	−2.24	−3.441	−3.650	−3.380
	(.017)	(.020)	(.021)	(.014)
N	126,000	126,000	126,000	126,000

NOTE.—Models 1 and 2 are logit models showing the bias resulting from assuming that dyads are independent.

tendency to be friends with your neighbor ($B_2 = 1$) but no endogenous effect as a result of mutual friends. Although model 1 for scenario 1 shows that the bivariate result is overestimated ($\hat{B}_1 = .714$), this is easily remedied in model 2 which includes the other covariate, residential proximity, in the logit model ($\hat{B}_1 = .571$).

In scenario 2, there is the same effect of racial homophily, plus an effect of mutual friends ($B_3 = .6$), but no effect of residential proximity. In this case, even though race is the only exogenous variable in the model, model 1 overestimates the homophily preference ($\hat{B}_1 = .713$). As suggested above, however, one might argue that the bivariate result is the "correct" measure because it incorporates both the direct and indirect effect of race-related preferences because of the fact that race is the only exogenous variable in the model.

Nevertheless, things get problematic in scenario 3 when we try to parcel out the effect of race per se from other factors that may affect friendship formation. Scenario 3 combines the variables from scenarios 1 and 2 to show what happens when we have two independent variables and an endogenous effect because of mutual friends. Model 1 of scenario 3 shows that the effect of race is significantly overestimated when it is the only explanatory factor. More important, even when spatial proximity, the other independent variable, is included in model 2, the effect of race is still overestimated by 67%. Clearly, in this constructed example, the bias is the result of the endogenous effect of mutual friends. Although all dyads begin the simulation with no friends in common, the final results indicate that same-race dyads have 2.5 mutual friends in common, compared to 1.7 for different race dyads. The important conclusion of model 2 of scenario 3 is that the effect of racial homophily is upwardly biased even when all the exogenous covariates are included in the model.

Scenario 4 of table 2 depicts an additional type of network endogeneity. Instead of mutual friends, it is possible that having a friend who lives nearby a potential friend would make a friendship more likely. In other words, playing or hanging out in your friend's neighborhood puts you into contact with other young people in that neighborhood and makes you more likely to be friends with them. Call this a "second-order" proximity effect. Model 1 of scenario 3 shows that the effect of race is again overestimated in a bivariate regression, and model 2 confirms that this holds even when the proximity effect is included.

The results of the analysis of simulated network data in models 1 and 2 of table 2 show that conventional logit models may substantially overestimate the effect of race when there is dependency in friendship choices of different individuals and when there are other factors that are correlated with race such as residential propinquity that also affect friendship. While real-life networks are likely to be more complicated than the data generated in table 2, the basic point should be clear: if we want to control for other factors that may contribute to racial friendship segregation, such as residential segregation or social class, conventional statistical models may lead to substantially biased estimates of the net effect of race, even when we have perfect information on the other variables in the model. In appendix A we discuss statistical models that attempt to account for the problem posed by the endogeneity of network structure and incorporate the key insights of this discussion into our analytic strategy, described below.

DATA AND METHODS

Add Health is a longitudinal school-based study of adolescents. It began in 1994/95 as a survey of young people in grades seven through twelve in 134 public and private schools (Bearman, Jones, and Udry 1997). The first wave of data collection included "in-school" questionnaires for all students present the day of the interview, our source of data on friendships; and "in-home" interviews with a subset of these students, our source of information about residential location. The in-school questionnaire asked the students to list their top five male and top five female friends in order of preference.[7] Data on the characteristics of nominated friends are obtained by matching the identification numbers of friendship nominations to respondents' identification numbers. Residential proximity is one of these characteristics.

In this article, the dyad is the unit of analysis. We study friendships and potential friendships among the 14,500 adolescents for whom information about the spatial location of their home in relation to the school (or other central place) was also collected. Each case consists of a dyad of two students who attend the same school, and the dependent variable is whether or not the students were friends. We consider all possible combinations $\{X_{i \to j} : i = 1, \ldots, N, j = 1, \ldots, N, i \neq j\}$ of the N students from the same school selected for the in-home interview. This results in $N\,(N - 1)$ cases per school. The dyad $X_{i \to j} = 1$ if person i nominated person j as a friend and zero otherwise. Note that the dyads are directional. There is no presumption that if person i nominated person j as a friend, then person j also nominated person i as a friend. Including all possible dyads of potential friends results in a data set of 3,878,011 cases, including 20,911 friendship dyads.

Modeling friendship networks is complicated because, as shown in the

[7] Requesting separate lists for friends of each gender may reduce the salience of gender in this study compared to others that used a less directive approach to data collection (cf. Shrum et al. 1988).

earlier example, friendship choices are influenced by the choices that others make. As a result, the standard statistical assumption of independence does not hold. Appendix A reviews the issues and discusses two approaches to estimation, the conventional logit model and the p^* model. The p^* model is the preferred model, as table 2 suggests it does a significantly better job estimating racial homophily than the conventional logit model, but as explained in appendix A, caution should be exercised in interpreting the results. Below we discuss the difference between the estimated homophily coefficients between the p^* and conventional logit models at length and argue that part of the difference between them is that the logit model estimates direct and indirect (via network endogeneity) effects of race on friendship while the p^* model attempts to isolate the direct effect of racial homophily.

Following Moody (2001), we use racial heterogeneity, h, as our measure of school-level racial diversity. Racial heterogeneity, or h, is the probability that any two randomly chosen students will be of a different race/ethnicity, $h = 1 - \sum_r p_r^2$, where p_r is the proportion of students in the school who are of race or ethnic group r. The measure h is a succinct way to measure racial diversity, and because, as discussed above, the dyad-based structure of our data already captures the role of propinquity (e.g., the dyads include all potential friends in the school), h measures "school climate" effects above and beyond the basic numerical effect of school composition. There are other possible specifications of school-level racial composition, but because there are only 134 schools in the data and limits on the variability in school composition for the different racial groups, the number of possible school-level composition variables, and their explanatory power, is constrained.[8] As a supplement to our main results, we also estimated models of school composition effects using the log of h to estimate a curvilinear effect of school heterogeneity (see n. 16 below).

In order to correctly estimate the effect of school composition effects on friendship formation, this article uses hierarchical models to estimate the conventional logit and p^* models (see Raudenbush and Bryk 2002). The data used in this article consist of 3.8 million dyads of potential friends nested in 134 schools, so there is much more variation at the dyad level than at the school level. We estimate the following model in HLM. The intercept and the coefficients on race and the network (p^*) variables

[8] For example, Quillian and Campbell (2003, pp. 561–62) show that there are only a handful of schools in the Add Health data in which blacks, Hispanics, or Asians are a numerical majority.

are estimated as random coefficients that are allowed to vary across schools. Equation (2) is the individual (level 1) model:

$$y_{ijk} = \beta_{0k} + \beta_{abk}\mathrm{Race}_{ij} + \beta_{p\mathrm{star}k} X_{p\mathrm{star}ij} + \alpha X_{ij}. \tag{2}$$

The dependent variable, logit y_{ijk}, indicates whether the ij dyad in school k are friends. β_{0k} is the intercept for school k, β_{abk} represents the coefficient on racial homophily for $a \rightarrow b$ dyads (where a is i's race and b is j's race) in school k, $\beta_{p\mathrm{star}k}$ represents a vector of coefficients on the network variables for school k (see app. A for a full description of these variables), and α represents a vector of coefficients for other explanatory variables. The coefficients β_{0k}, β_{abk}, and $\beta_{p\mathrm{star}k}$ are allowed to vary across schools according to the school-level equations (3)–(5):

$$\beta_{0k} = \beta_0 + \delta_1 h_k + \mu_{0k}, \tag{3}$$

$$\beta_{abk} = \beta_{ab} + \delta_{ab} h_k + \mu_{abk}, \tag{4}$$

$$\beta_{p\mathrm{star}k} = \beta_{p\mathrm{star}} + \delta_2 \log(\mathrm{pop})_k + \mu_{2k}, \tag{5}$$

where μ_{0j}, μ_{abj}, and μ_{2j} are normally distributed error terms with mean zero and variance σ_0^2, σ_{ab}^2, and σ_2^2. School racial heterogeneity is included as an explanatory variable for variation in the intercept and the coefficients on racial homophily across schools, and the natural log of the number of students in the school is included as an explanatory for variation in the p^* coefficients.

SOCIAL AND RESIDENTIAL SEGREGATION: DESCRIPTIVE RESULTS

Table 3 shows the observed level of racial segregation among the friendship dyads in the data. Rows indicate the race of the respondent and the columns indicate the race of the nominated friend. We focus on four race and ethnic groups: whites, blacks, Asians, and Hispanics. In this article, respondents who indicate a Hispanic or Latino origin are coded as Hispanic regardless of how they respond to other questions about their race or ethnicity. "Other" refers to Native Americans, who are too few in number to support a separate category, and also to cases where information on race and ethnicity is missing.

Not surprisingly, students tend to have friends of the same race and ethnic group, as shown by the large number of cases on the diagonal. However, the level of social segregation seems to vary by group. Whereas 85.1% of white students' nominated friends are white, and 71.4% of black students' friends are black, only 42.3% and 33.5% of the friends of Asian

TABLE 3
Racial Composition of Friends, by Race

Respondent's Race	Friend's Race (Percentage)						
	White	Black	Asian	Hispanic	Other	Total	N
White	85.1	1.4	1.1	4.8	7.6	100	10,835
Black	10.7	71.4	1.6	9.8	6.6	100	3,389
Asian	35.8	8.8	42.3	7.1	6.0	100	1,889
Hispanic	40.9	12.4	1.7	33.5	11.6	100	3,191
Other	68.1	6.9	.8	14.3	10.0	100	1,557
Total	73.2	8.6	2.0	8.3	8.0	100	20,861

Note.—Table 1 contains only friendship dyads, $D_{ij} = 1$.

TABLE 4
Race of Other Students Living within One-Half Kilometer of the Respondent

Respondent's Race	White	Black	Asian	Hispanic	Other	Total	Unweighted N
White	81.5	2.3	6.2	4.9	5.2	100	2,092
Black	14.5	60.6	2.4	11.6	10.9	100	765
Asian	20.6	2.6	45.7	26.2	5.0	100	436
Hispanic	40.2	9.8	10.1	31.6	8.4	100	864
Other	69.2	9.2	5.4	4.5	11.6	100	238
Total	66.7	9.4	8.4	9.2	6.4	100	4,395

and Hispanic students are of the same race. This might reflect preferences, but it might also represent differences in opportunity, given that the representation of Asian and Hispanic groups is much lower. Looking at cross-group nominations, the effect of group size is immediately evident. The proportion of whites nominating friends of another race is much lower than the proportion of those of other races nominating white friends. White students outnumber others by a factor of 3-1 or more in the Add Health data.

Table 4 provides a simple descriptive overview of the level of residential segregation in the schools in the Add Health data by indicating the race of other students—both friends and nonfriends—who live within .5 km of the respondent. This is a relative measure because there are no fixed boundaries for neighborhoods—rather, the notion of a neighborhood is viewed as relative to the students in question. In table 4, residential segregation is evident among all of the race/ethnic groups. Students are more likely to live near students of their own race than would be the case if neighborhoods were perfectly integrated. For example, 81.5% of the students within .5 km of white students are white. Overall, the similarities between tables 3 and 4 reinforce the expectation that residential segregation has the potential to affect patterns of social segregation in schools. However, there may be other reasons for the similarity. To interpret patterns of social segregation, it is important to consider dyads of potential friends, to see whether students' residential location affects whom they choose as their friends.

Table 5 presents summary statistics of variables used in the analysis for friend and nonfriend dyads. Comparing the means of this table provides a rough indication of the variables that affect friendship probabilities. Friends are more likely than nonfriends to be of the same race or ethnicity. Compared to nonfriends, friends are also more likely to be of the same grade, socioeconomic background, and gender. Clearly, grade in school is an important factor, related both to preferences but also to propinquity (i.e., greater likelihood of shared classes). With respect to socioeconomic status (SES), we include variables measuring the absolute value of the difference in parents' education (years of schooling), and income (thousands of dollars). The absolute difference in parents' income, occupation, and most noticeably education are smaller for friend dyads than for nonfriend dyads. Because parental information is less available in the Add Health data than other kinds of information, and we did not want to restrict our sample, we created separate variables to indicate missing data for the parents of one or both of the potential friends in the dyad.

Various measures of residential location are presented in table 5. The in-home sample of Add Health contains the (relative) latitude and longitude of the student's home, allowing us to calculate the Euclidean distance between the residences of dyad members. The simplest of the residential proximity measures divides this distance into four categories, thus allowing a flexible functional form specification for the effect. Table 5 suggests that distance matters, as friends are more than five times as likely as nonfriends to live within .25 km of one another. The difference narrows from 5.2 to 3.0 if we expand the distance to .50 km (inclusive), and narrows further to 1.88 if we expand to 1 km. We consider two alternative functional form specifications for the distance. "Log of distance" is the natural log of the distance between the two houses in a dyad. To take into account differences between the sizes of school catchment areas in the examination of residential proximity, we also created a measure of called "relative weighted distance," which is calculated as

$$\mathrm{RWD}_{ij} = \frac{e^{-.2(\mathrm{dist}_{ij})}}{(1/N)\sum_{j=1}^{N} e^{-.2(\mathrm{dist}_{ij})}},$$

where dist_{ij} is the distance between students i and j, and N is the number

TABLE 5
Summary Statistics for Friend and Nonfriend Dyads

Variable	Nonfriend Dyads Mean	Nonfriend Dyads SD	Friend Dyads Mean	Friend Dyads SD
Dyad type:				
White→white	.200		.432	
White→black	.027		.013	
White→Asian	.015		.009	
White→Hispanic	.027		.028	
Black→white	.027		.011	
Black→black	.064		.121	
Black→Asian	.033		.004	
Black→Hispanic	.046		.015	
Asian→white	.015		.009	
Asian→black	.033		.004	
Asian→Asian	.053		.065	
Asian→Hispanic	.057		.009	
Hispanic→white	.027		.025	
Hispanic→black	.046		.015	
Hispanic→Asian	.057		.009	
Hispanic→Hispanic	.138		.096	
Other→other	.134		.134	
Same grade	.284		.745	
Same sex	.495		.602	
Difference in:				
Parents' education	2.086	2.619	1.774	2.067
Parents' income	1.386	1.640	1.608	1.628
Missing:				
Parents' education	.304		.215	
Parents' income	.532		.443	
Clubs and sports:				
No. of school clubs in common	.074	.317	.286	.611
No. of school sports in common	.102	.365	.367	.690
Distance measures:				
Relative weighted distance	.999	.280	1.171	.841
Ln(distance)	8.154	1.067	7.873	1.710
Physical distance:				
>2 km	.729		.657	
1–2 km	.167	.373	.148	.355
.5–1 km	.068	.252	.089	.284
.25–1/2 km	.024	.152	.044	.206
<.25 km	.012	.110	.062	.241
Have friend within .25 km	.068		.208	
Network variables:				
Reciprocity	.003	.056	.400	.490

TABLE 5 (*Continued*)

Variable	Nonfriend Dyads Mean	Nonfriend Dyads SD	Friend Dyads Mean	Friend Dyads SD
Transitivity	.074	.465	3.590	3.801
Expansiveness	1.953	2.355	3.244	3.205
Popularity	1.955	2.228	2.832	2.282
N	3,857,712		19,959	

of students in the school. The numerator is a distance weight that gets smaller as the distance between the potential friends gets larger.[9]

Finally, the social network variables used in the p^* model are depicted at the bottom of table 5. These are calculated according to equations (A4)–(A8) in appendix A as the change in the network statistics when i and j are friends versus when they are not friends. These variables were constructed using the full in-school sample of all the friends in the data. Table 5 shows that 40.0% of friend dyads ($x_{ij} = 1$) are reciprocated (i.e., $x_{ji} = 1$), while in nonfriend dyads, j nominates i as a friend only .3% of the time. In addition, friend dyads have on average 3.59 mutual friends in common versus only .074 for nonfriend dyads.

DETERMINANTS OF FRIENDSHIPS IN SCHOOL: RESULTS

Table 6 presents several models of the likelihood of friendship among dyad pairs. Of primary interest are the coefficients for the different race and ethnic categories. The excluded category is white → white, so the coefficients indicate the relative likelihood of friendship among the other categories. The dyad racial categories are estimated as random coefficients, thereby allowing the effect of race to vary across schools. Model 1 presents results for a model of the effect of race controlling for grade and gender. Table 6 gives the point estimates for the coefficients, and table 7 gives the variance for the random effects. In model 1 of table 6, all the interracial combinations have negative coefficients, meaning that the likelihood of these friendships is less than white → white friendship, controlling for grade and gender. For example, the coefficient on white → black dyads is -1.318 (SE = 136; $P < .001$) relative to white → white dyads, while the coefficient on white → Asian dyads is $-.796$ (SE = 121; $P < .001$).

Model 2 of table 6 adds school racial diversity, h, as a school-level variable to explain differences in the coefficients on race across schools.

[9] For relative weighted distance, $-.2$ is used because it was the coefficient on distance in a one-variable logit model of friendship.

TABLE 6
LOG ODDS OF FRIENDSHIP REGRESSED ON EXPLANATORY VARIABLES

	Model 1		Model 2		Model 3		Model 4		Model 5	
Dyad Type[a]	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
White→black	−1.318	.136**	−1.281	.126**	−1.229	.124**	−1.208	.126**	−.640	.105**
School diversity			−3.100	.726**	−2.974	.716**	−2.928	.726**	−1.329	.593*
White→Asian	−.796	.121**	−.782	.109**	−.728	.108**	−.673	.109**	−.305	.114**
School diversity			−1.939	.566**	−1.809	.565**	−1.788	.568**	−1.213	.584*
White→Hispanic	−.549	.083**	−.512	.068**	−.479	.067**	−.439	.067**	−.134	.064*
School diversity			−2.666	.372**	−2.553	.368**	−2.489	.367**	−1.451	.329**
Black→white	−1.451	.144**	−1.372	.122**	−1.352	.125**	−1.311	.122**	−.753	.100**
School diversity			−3.946	.718**	−3.880	.730**	−3.780	.716**	−2.566	.581**
Black→black	.532	.116**	.506	.120**	.430	.113**	.460	.116**	.373	.079**
School diversity			−2.150	.686**	−2.111	.638**	−2.134	.657**	−.269	.422
Black→Asian	−1.001	.200**	−.672	.231**	−.754	.238**	−.642	.230**	−.468	.252
School diversity			−5.646	1.138**	−5.459	1.180**	−5.330	1.133**	−2.826	1.166*
Black→Hispanic	−.402	.123**	−.427	.120**	−.424	.124**	−.380	.121**	−.039	.102
School diversity			−2.728	.638**	−2.569	.661**	−2.507	.641**	−.786	.499
Asian→white	−.886	.118**	−.889	.110**	−.837	.106**	−.799	.110**	−.415	.111**
School diversity			−2.297	.580**	−2.126	.564**	−2.150	.579**	−1.810	.549**
Asian→black	−.714	.213**	−.396	.225	−.445	.230	−.349	.223	−.025	.240
School diversity			−5.627	1.155**	−5.513	1.175**	−5.334	1.138**	−2.444	1.186*
Asian→Asian	.587	.176**	.775	.222**	.706	.205**	.808	.215**	.740	.211**
School diversity			−1.972	1.039	−2.156	.933*	−2.164	.988*	−3.397	.909**
Asian→Hispanic	−.891	.151**	−.677	.177**	−.670	.186**	−.540	.175**	−.307	.205
School diversity			−3.326	.822**	−3.501	.862**	−3.551	.808**	−1.621	.866
Hispanic→white	−.715	.090**	−.683	.077**	−.648	.076**	−.606	.075**	−.313	.066**
School diversity			−2.425	.422**	−2.294	.414**	−2.235	.411**	−.911	.339**
Hispanic→black	−.359	.141**	−.358	.137**	−.357	.140*	−.333	.136*	−.058	.127
School diversity			−3.467	.735**	−3.204	.753**	−3.176	.728**	−.783	.617
Hispanic→Asian	−1.133	.129**	−1.034	.205**	−.966	.212**	−.899	.204**	−.852	.257**

School diversity			−2.142	.908*	−2.488	.939**	−2.328	.899*	.555	1.096
Hispanic→Hispanic	−.068	.098	−.138	.110	−.164	.110	−.114	.111	.118	.092
School diversity			−1.210	.564*	−1.063	.566	−1.044	.573	−.208	.425
Other→other	−.528	.065**	−.513	.053**	−.495	.053**	−.455	.053**	−.083	.041*
School diversity			−1.970	.295**	−1.902	.297**	−1.910	.293**	−.895	.206**
Same sex	.439	.015**	.439	.015**	.440	.015**	.438	.015**	.287	.018**
Same grade	2.187	.017**	2.188	.017**	2.201	.017**	2.200	.017**	1.301	.019**
Other friend <.25 km					.801	.020**	.798	.020**	.315	.025**
Physical distance:										
.5–1 km					.314	.027**	.310	.027**	.199	.032**
.25–.5 km					.539	.037**	.533	.037**	.394	.045**
Less than .25 km					1.217	.033**	1.209	.033**	.762	.042**
Difference in:										
Parents' education							−.055	.004**	−.026	.005**
Parents' income							−.031	.009**	−.034	.011**
Missing:										
Parents' education							−.370	.021**	−.095	.025**
Parents' income							−.115	.031**	−.128	.037**
Network variables:										
Reciprocity									2.786	.068**
Log of no. of students									.184	.051**
Transitivity									.573	.015**
Log of no. of students									.027	.012*
Expansiveness									.105	.009**
Log of no. of students									−.048	.006**
Popularity									.073	.021**
Log of no. of students									−.141	.010**
Constant	−5.781	.095**	−5.786	.095**	−5.960	.089**	−5.681	.093**	−6.244	.084**
H, G01			−.347	.523	−.301	.487	−.264	.486	.531	.183**
N	3,857,712		3,857,712		3,857,712		3,857,712		3,857,712	

[a] Excluded category: white→white.

* $P < .05$.

** $P < .01$.

TABLE 7
VARIANCE COMPONENTS FOR RANDOM EFFECTS

Random Effect	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	.985**	.988**	.913**	.911**	.741**
White→black	1.083**	.865**	.841**	.861**	.564
White→Asian	.727**	.570**	.539**	.559*	.496*
White→Hispanic	.664**	.437*	.426*	.421**	.287*
Black→white	1.144**	.810**	.825**	.806**	.492*
Black→black	1.039**	.940**	.867**	.897**	.523**
Black→Asian	1.245**	.903**	.966*	.871**	.579*
Black→Hispanic	.974**	.740**	.767**	.753**	.475**
Asian→white	.711**	.549*	.509*	.550*	.525
Asian→black	1.366**	1.024**	1.049**	.999**	.781**
Asian→Asian	1.010**	.920**	.733**	.770**	.714
Asian→Hispanic	.867**	.589**	.604**	.525*	.689**
Hispanic→white	.721**	.528**	.507**	.499**	.241*
Hispanic→black	1.156**	.902**	.925**	.892**	.739**
Hispanic→Asian	.665**	.473	.478	.397	.584
Hispanic→Hispanic	.723**	.655**	.679**	.695**	.417
Other→other	.572**	.422**	.427**	.416**	.234
Reciprocity					.658**
Transitivity					.152**
Popularity					.076**
Expansiveness					.206**

* $P < .05$.
** $P < .01$.

Each of the "school diversity" variables in models 2–5 should be thought of as an interaction term affecting the coefficient on racial homophily. A negative coefficient on these school diversity interaction terms means that increasing school diversity reduces the relative probability of friendship for that type of racial dyad. The variable h is centered at .46, which is the average level of racial diversity in the Add Health schools (the range is .2–.76). As a result, we can use the coefficients in models 2–5 to predict the degree of racial homophily in schools with different levels of diversity. For example, the log odds of a white → black dyad being friends relative to a white → white dyad in a school with an h of .56 would be $-1.281 + (.56 - .46) \times (-3.100) = -1.591$. Because the coefficient on school diversity is negative for all the interracial dyads, model 2 predicts that the probability of interracial dyads being friends decreases in more diverse schools. Thus, although more diverse schools have, overall, more potential interracial contact and hence more interracial dyads of "potential" friends, the probability that any one of them is a friend dyad is smaller relative to less diverse schools. This corroborates the results of Moody (2001), extending them for each of the possible interracial combinations. Later in the article, we will use the coefficients from models 1 and 2 to assess the magnitude of the effect of school diversity on interracial friendship.

Model 3 of table 6 adds measures of the spatial distance between potential friends. We estimate the effect of distance using a categorical variable for the distance between members of the dyad that distinguishes distances (within .25 km, .25–.49 km, .50–1.0 km, and more than 1 km). The default category is a pair that lives more than 1 km apart. Does residential proximity affect friendship formation? The answer is yes. You are more likely to be friends with someone if you live between .5–1 km ($b = .314$; $P < .01$) or .25–.49 km ($\beta = .539$; $P < .01$) from him or her. The effect is substantially larger for individuals who live within .25 km of each other ($\beta = 1.217$; $P < .01$). A distance of .25 km is approximately equal to the length of two football fields (including the endzones). We call this large effect of dyads who live close to each other the "bus stop effect" to emphasize its local nature.

In addition to the distance between pairs of the potential friends, model 3 also adds a variable indicating whether one of the dyad members had a friend that lived within .25 km of the other dyad member. This is to control for the possibility that there is an indirect effect of spatial proximity. If you visit friends who live in a different neighborhood than your own, it is possible that you may be more likely to become friends with other students in that neighborhood. As discussed above in reference to table 2, we hypothesize that this is one mechanism that may magnify the effects of residential segregation. The coefficient on this variable is large and positive (.801; $P < .01$), indicating that you are significantly more likely to be friends with someone if you have friends who live near that person. We call this the "indirect bus stop effect."

So far, we have shown evidence of a direct and indirect effect of spatial proximity on friendship. We have also shown evidence of residential segregation: students living closer to each other are more likely to be of the same race. Does residential segregation within school districts lead to friendship segregation within schools? By comparing the coefficients on race in models 2 and 3 we find that the answer is yes, but not very much. Adding the variables for distance between dyad members as well as the effect of nearby friends reduces the coefficients on the race and ethnic categories, but the reduction is small. For example, in model 2, the coefficient for white → black dyads is -1.281; this drops to -1.229 in model 3 when we add the distance measures (a decline of 4.1%). For black → white dyads it declines from -1.372 to -1.352.[10] What model 3 suggests

[10] Noting that the predicted log odds of a black → white versus a black → black friendship is $\beta_{\text{black}\rightarrow\text{white}} - \beta_{\text{black}\rightarrow\text{black}}$, the log odds of black → white versus black → black

is that even if residential segregation were completely eliminated *within* school districts, friendship segregation would decline only slightly. This result is not an artifact of the use of a categorical variable to measure the distance between potential friends. Nor is it confounded by differences in the size of school catchment areas. In appendix B, we show that our results are robust to the alternative measures of distance discussed above, the log of distance and relative weighted distance.

In model 4, we control for social class differences by adding measures of the difference in parents' income and education. Both of these variables have statistically significant effects on friendship formation. Increasing the difference in parents' years of education, for example, decreases the probability of friendship ($\beta = -0.055$; $P < .001$). However, including the social class variables has little effect on the racial homophily coefficients; the coefficient on white → black friendship, for example, declines from -1.229 to -1.208. Similar to residential proximity, social class has a small and measurable effect on school friendship, but it does not explain much of the racial segregation of friendship.

Model 5 of table 6 presents results for the p^* model. The p^* model attempts to take the dependent nature of the data into account—that is, the fact that potential dyads are not independent observations. The p^* model adds the network variables (reciprocity), the number of mutual friends (transitivity), the number of friends i nominated (expansiveness), and the number of people who nominated j as a friend (popularity). These were measured for each school from the full in-school sample. The effects of the network variables are estimated as random coefficients, which allows them to vary across schools. The log of the number of students in the school (in the full in-school sample, standardized to the mean) is included as a predictor of the coefficients of the social network variables (see eq. [13]), under the assumption that network effects will be larger in big schools. Not surprisingly, the social network variables all have strong effects on the likelihood of friendship. Because the estimation strategy is pseudo likelihood, the standard errors are only approximations (Strauss and Ikeda 1990). The results in model 5 indicate that you are much more likely to be friends with someone if that person is also friends with you (2.786; $P < .001$) or if you share mutual friends (.573 per mutual friend; $P < .001$).

What is striking about model 5 is that all of the coefficients on the race and ethnic categories decline substantially from model 4. The magnitude of the coefficient for white → black dyads in model 5, for example, declines by 47% compared to model 4. As discussed above, the p^* model attempts to control for the endogeneity of network structure, the fact that one's

friendship declines from $-1.872(-1.372 - .506)$ to $-1.782(-1.352 - .430)$.

friendship choices are affected by the friendship choices of other people in the data. It is useful to think of the way the friendship choices of your friends affects your choices in terms of "network propinquity." Friendship between any two individuals who go to the same school does not occur at random, but is constrained by the likelihood that they will interact with each other, and this in turn is affected by the network of friendship ties that connects them together. The degree of network propinquity is evident in the data; because whites are more likely to be friends with each other than with blacks, for example, part of the effect of race for whites in models 1–4 is being attributed to greater network proximity—the degree of reciprocity and the number of mutual friends—in the p^* analysis in model 5. Among friendship dyads, 44% of white-white friendships are reciprocated compared to 27% of white-black friendships, and the number of mutual friends in common (transitivity) is 5.8 for white-white and 2.6 for white-black friend dyads (results available on request). In other words, the p^* model attributes some of the effect of race in models 1–4 to network propinquity in model 5.

When evaluating the difference between the coefficients in the independence (models 1–4) and p^* models (model 5) it is important to consider the different sources of network propinquity. Part of the network effect is the result of the indirect effect of race, and part of it is the result of the effect of other factors that affect friendship. In the simulation results presented earlier in table 2, the indirect effect of race is evident in scenario 2. As discussed above, in scenario 2 race is the only factor in the model, so the difference between the results in model 1 of table 2 (which estimates a logit model assuming dyad independence) and model 3 of table 2 (which estimates a p^* model) depends on what you want to estimate; the underlying homophily parameter ($\beta = .6$ in scenario 2) or the resulting level of segregation ($\hat{\beta} = .713$). Hence, part of the difference in the coefficients on race in models 4 and 5 of table 5 is attributable to the indirect effect of race. Because of a concern that network variables such as mutual friends may be "explaining away" the effect of race on friendship choice, Quillian and Campbell (2003), for instance, only include the number of friends nominated and the respondent's popularity as controls for network structure.[11]

On the other hand, network propinquity may also incorporate the indirect effect of other factors that are correlated with race. The simulation results in table 2 show that if there are other factors that affect friendship

[11] Because the number of friends nominated and popularity are individual rather than dyad-specific variables they only affect the marginals (i.e., the total number of friends that each person chooses) rather than the relative probability of choosing friends of particular race and ethnic categories (see Holland and Leinhardt 1981, p. 36).

that are correlated with race, such as residential segregation or social class, then the effect of race is likely to be overestimated in statistical models that assume independence. Thus, as discussed above, the estimated coefficient on race in the simulation of scenario 3 of table 2 is overestimated by 67% in model 2 (which includes the exogenous covariates but not the network variables). This demonstrates that the effect of race per se may be biased if the effect of network propinquity is not controlled for. Therefore, in the presence of other factors that affect friendship and are associated with race, such as social class and residential location, it is incorrect to infer the effect of race from logit models that assume independence.

This raises an important question. What outcome level of racial segregation would the p^* model predict in the absence of all other factors except the direct and indirect effects of race and ethnicity? In other words, how much racial friendship segregation would result in our data if everyone were identical except for race, and where the preference for racial homophily and the tendency for reciprocity and transitivity were the estimated coefficients in model 5 of table 6? Conceptually, this is simply a matter of predicting the outcome variable holding everything except race and the network variables constant. However, because the network variables are endogenous, we have to simulate the data instead, similar to the procedure in table 2. Because such a simulation is prohibitive for 3.8 million cases and 134 schools, we opt instead to simulate the results for a hypothetical school of 100 students, 50% of race A and 50% of race B, resulting in 9,900 possible friendship dyads. We simulated the "predicted" level of white-black segregation using the coefficients on white → black dyads, reciprocity, and transitivity in model 5 of table 6.

Appendix C provides details on the process of simulating the overall level of segregation based on the results in table 6. Table C1 shows that the *predicted* level of white → black friendship segregation based on 20 simulations was −.976, which is substantially larger than the estimated p^* white → black homophily coefficient of −.64 in model 5. The difference between the predicted value of −.976 and the model 5 homophily coefficient of −.64 reflects the indirect effect of race operating through network propinquity, that is, the tendency of one's friends to choose friends of the same race. In addition, the predicted level of −.976 is 19% smaller than the estimate in model 4, −1.208, suggesting that the overall level of white-black segregation is overestimated in models that assume dyad independence.[12] Again, this result is not surprising when you recall the findings

[12] Because the effect of reciprocity and mutual friends is the same for all groups in table 6, we expect that predicting the segregation levels for the other race and ethnic dyads would reveal a similar level of overestimation.

from scenario 3 of table 2: if there are other factors that affect interracial friendship, then estimates of racial homophily will be upwardly biased in statistical models that assume dyad independence even when the other factors are included as covariates in the model. At the same time, these results should be interpreted cautiously, as they depend on both the estimates of the p^* model, which, as mentioned above, is pseudo-maximum likelihood, and the results of simulated friendship data based on the assumption that the effects of network structure have been identified correctly. Checking the predicted results in table C1 with more elaborate models of network structure should be a subject of future research.

CALCULATING THE EFFECT OF RESIDENTIAL SEGREGATION ACROSS SCHOOL DISTRICTS

Although the residential proximity variables explained very little of the observed level of friendship segregation within school districts, we have yet to consider the effect of residential segregation across them. Obviously, if students go to local schools, then residential segregation across school districts restricts the opportunity for individuals to become friends. In this section we calculate the effect that residential segregation across school districts has on in-school friendship segregation.

Based on the estimated homophily coefficients in table 6, we can calculate the proportion of the overall level of friendship segregation that is attributable to the segregation of students across schools. If we assume that the preference for racial homophily is the same across schools (i.e., if we ignore, for the moment, the effect of school racial heterogeneity), then appendix D shows that we can express the log-odds ratio of interracial friendship segregation between races i and j into components based on school segregation and homophily preferences:

$$\mathrm{Seg}_{ij}^{A} = S_{ij} + \alpha_{ij} - \alpha_{ii} + \alpha_{ji} - \alpha_{jj}, \tag{6}$$

where S_{ij} is the log-odds ratio of school segregation between i and j across all the schools, and α_{ij}, α_{ii}, α_{ji}, and α_{jj} are homophily coefficients estimated in table 6.[13] The percentage of friendship segregation attributable to residential segregation is $100 \times (S_{ij}/\mathrm{Seg}_{ij})$.

The benefit of the approach depicted in equation (6) is that it provides a clear decomposition of the contribution residential segregation across school districts plays in friendship segregation. Nonetheless, it does not take the school-level effect of racial heterogeneity, h, into account. As

[13] The log odds of school segregation is calculated as $S_{ij} = \ln(s_{ij}s_{ij}/s_{ii}s_{jj})$, where s_{kl} is the average proportion of students of race k in schools attended by students of race l.

discussed above, h controls for "school climate" effects on the probability of interracial friendship above and beyond the basic propinquity effect (which is controlled for by using dyad data on potential friends). Although the number of potential interracial friends increases in racially heterogeneous schools, the coefficient on h in models 2–5 of table 6 indicates that the probability of any particular interracial dyad being friends goes down. Therefore, increasing diversity seems to result in larger preferences for racial homophily. Appendix D shows we can predict the overall odds ratio of friendship segregation between i and j incorporating the effect of h as

$$\mathrm{Seg}_{ij}^{B} = \ln \frac{\sum_k t_{ij}^k \sum_k t_{ji}^k}{\sum_k t_{ii}^k \sum_k t_{jj}^k}, \tag{7}$$

where t_{ij} is the predicted number of $i \rightarrow j$ friends in school k based on school composition and the homophily coefficients from table 6. If we calculate Seg_{ij}^{B} twice, once using the existing school enrollment data and then under school integration, we can calculate the proportion of the overall level of friendship segregation that is attributable to the segregation of students across schools.

Tables 8 and 9 present descriptive data on school segregation, and table 10 provides estimates of the overall effect of school segregation on friendship segregation. Table 8 depicts the level of school segregation by showing the average school composition, by race, in the Add Health data.[14] The level of school segregation between whites and blacks, S_{WB}, for example, can be calculated directly from tables 8 and 9 as $\ln[(.445 \times .135)/(.663 \times .319)] = -1.26$. However, because of the multistage aspect of the sample design, the level of school segregation in the Add Health data combines the effect of residential segregation within metropolitan areas with the effect of the uneven distribution of racial groups across cities and states in the United States. In order to calculate the level of within-metropolitan-area school segregation, we use data from the National Center for Educational Statistics (NCES) on the school composition of public middle schools and high schools in all 328 urban areas in the United States.[15] Table 9 shows the average school composition by race for the NCES data.

[14] School racial composition was calculated from the full in-school data.

[15] Although public schools typically draw students from local neighborhoods, it is possible that school segregation differs from the underlying level of residential segregation if large numbers of students go to magnet schools, and so on. Nonetheless, tract-level residential segregation in the 2000 census and school segregation in the NCES are highly correlated at the metro level ($r = .86$; authors' calculations), and this correlation would be higher if residential segregation were calculated by school zoning districts rather than census tracts (because there is tract-level segregation within school zoning districts).

TABLE 8
AVERAGE SCHOOL COMPOSITION, BY RACE: ADD HEALTH DATA

	Average Proportion in School				
Race	White	Black	Asian	Hispanic	Total
White	.663	.135	.053	.148	1
Black	.445	.319	.057	.180	1
Asian	.487	.157	.143	.213	1
Hispanic	.453	.167	.073	.308	1

TABLE 9
AVERAGE COMPOSITION, BY RACE: PUBLIC MIDDLE AND HIGH SCHOOLS IN U.S. URBAN AREAS, NCES DATA

	Average Proportion in School				
Race	White	Black	Asian	Hispanic	Total
White	.728	.128	.041	.103	1
Black	.447	.379	.041	.133	1
Asian	.483	.137	.169	.211	1
Hispanic	.361	.133	.063	.442	1

Using the data from the NCES on school segregation within U.S. metropolitan areas, table 10 estimates the effect of school segregation on friendship segregation; the top half uses equation (6), ignoring the effect of school diversity, and the bottom half uses equation (7), taking school diversity effects into account. The results in the top half of table 10 can be calculated directly from the coefficients in model 2 of table 6 and the school segregation levels in tables 8 and 9 using equation (6).[16] We estimate, for example, that 29% of white → black friendship segregation within U.S. urban areas is attributable to residential segregation within urban areas. In contrast, 42.2% of white → Hispanic segregation is attributed to school segregation. The overall effect of school segregation in the final column of table 10 is calculated by weighting the log odds according to each group's population size in the aggregate NCES data.[17] We estimate that residential segregation accounts for 32.8% of the overall level of school friendship segregation.

The bottom half of table 10 takes the effect of school racial heterogeneity

[16] We use model 2 of table 6, which only has controls for race, sex, and same grade because we are decomposing observed friendship segregation into within- and across-school components, not attempting to explain the within-school component (as in models 3–5).

[17] The weights are $(\text{pop}_i \times \text{pop}_j)/(\sum_{i \neq j} \text{pop}_i \times \text{pop}_j)$, where pop_i is the proportion of racial group i in all U.S. public middle schools and high schools.

TABLE 10

Predicted Percentage of Within-School Friendship Segregation Due to School Segregation in U.S. Urban Areas

	White→ Black	White→ Asian	White→ Hispanic	Black→ Asian	Black→ Hispanic	Asian→ Hispanic	Overall
No school heterogeneity effects:[a]							
School segregation S^{NCES} (all U.S. metro areas)	−1.37	−.70	−.97	−1.14	−.82	−.97	
Homophily preferences $(\alpha_{ij} - \alpha_{ii} + \alpha_{ji} - \alpha_{jj})$ from model 1 of table 6	−3.30	−2.28	−1.33	−3.36	−.85	−2.68	
Predicted log odds of interracial friendship $(A + B)$	−4.67	−2.98	−2.30	−4.50	−1.67	−3.65	
% friendship segregation due to school segregation $(100 \times [A/(A + B)])$	29.3%	23.4%	42.2%	25.3%	49.1%	26.6%	32.8%
With school heterogeneity effects:[b]							
Predicted segregation with current enrollment	−4.00	−2.81	−1.54	−3.10	−1.85	−3.20	(Average $H = .276$)
Predicted segregation without school segregation	−2.73	−2.23	−.72	−1.73	−.91	−2.16	(Average $H = .372$)
% friendship segregation due to school $100 \times [(C - D)/C]$	31.7%	19.9%	53.2%	44.1%	50.8%	32.6%	36.6%
NCES population weights	.36	.11	.36	.03	.10	.03	1.0

[a] Based on results in model 1 of table 6 and equation (6).

[b] Based on results in model 2 of table 6 and equation (7).

into account using equation (7) and the estimated coefficients of model 2 of table 6. The percentage of friendship segregation attributable to residential segregation is calculated as

$$\text{Percentage decline} = 100 \times \frac{\text{Seg}_{ij} - \text{Seg}_{ij}^{nrs}}{\text{Seg}_{ij}}, \tag{8}$$

where Seg_{ij}^{nrs} is the level of friendship segregation when school composition is equalized across schools within metropolitan areas. The results at the bottom of table 10 indicate that taking school diversity into account does not substantially alter the results from the top of table 10. The percentage of white-black friendship segregation attributable to residential segregation is 31.7%, and the overall effect is estimated to be 36.6%.[18]

In order to make sense of the results in the bottom half of table 10, there are two factors to consider. First, as discussed above, increasing h decreases the probability of interracial friendship. This decreases the numerator of equation (8). Second, however, the overall level of racial diversity in the Add Health data ($h = .46$) is higher than it is in public middle and high schools in the United States as a whole according to the NCES data ($h = .26$). This decreases the denominator of equation (8). These two effects work in opposite directions, and in the NCES data, they cancel each other out.[19]

Overall, the results in table 10 show that although we found little effect of spatial proximity within schools on racial friendship segregation, residential segregation across schools has a substantial effect. The results in table 10 suggest that about 30% of black-white and between 33%–37% of overall racial friendship segregation in schools is attributable to residential segregation across schools in U.S. metropolitan areas.

[18] In addition, we also estimated models using the log(h) instead of h to allow for a curvilinear effect of school heterogeneity. This resulted in a slightly higher level of friendship segregation attributable to residential segregation: 34.6% of white-black, 65.5% of white-Hispanic, and 45.0% overall (full results available upon request).

[19] An example: if a school system was composed of two schools, school A with 100 whites, 20 blacks, one Asian, and one Hispanic, and school B with 100 blacks, 20 whites, one Asian, and one Hispanic, then h would be .310 for each school, and .516 if they were aggregated together, eliminating segregation at the school level. Based on the coefficients in model 2 of table 6, Seg_{12} would decline from -4.29 to -3.43, a 20% decline. If there were no school heterogeneity effects, the decline would be from -5.07 to -3.16, a 37.6% decline. If, however, the school heterogeneity was .2 in both schools and .3 in the combined school (mirroring the increase in heterogeneity observed in table 10 with the NCES data), then the decline would be from -3.79 to -2.38, a 37.4% decline.

CONCLUSION

This article uses unique data from Add Health to study the effect of residential segregation on friendship segregation among middle and high school students. It attempts to overcome a methodological problem that has affected the previous literature on school friendship segregation: the endogeneity of network structure. As we show using simulated data in table 2, endogenous network effects, such as the role of mutual friends on friendship formation, may bias the results of statistical methods that assume the cases are independent. We estimate both conventional independence-logit models and p^* models, which attempt to control for social network structure. The p^* results suggest that the independence-logit estimates of racial homophily are upwardly biased by about 18% in these data. Again, however, we caution that the p^* model is pseudo maximum likelihood because it estimates the log odds of friendship for a dyad of potential friends conditional on the rest of the network. More research is needed to verify the accuracy of p^* estimates and develop feasible alternative methods.

The primary focus of this article is the effect of residential segregation on friendship segregation. The in-home sample of the Add Health data is ideal for this purpose because we have data on the latitude and longitude of the respondent's house with respect to his or her school or other focal point. This allows us to calculate precise measures of the distance between pairs of potential friends. Using several different measures of distance, we find that spatial proximity does affect friendship formation. However, the effect is very local; we call it a "bus stop effect" because most of the effect occurs for individuals who live within .25 km of each other. We also find evidence of an indirect effect of distance: the results in table 6 indicate that you are more likely to be friends with someone if you have a friend who lives within .25 km of him or her. This indirect effect of distance is sizable and significant, which suggests that the effect of space on friendship formation is more complicated than a simple function of the distance separating two individuals, but instead interacts with network structure to affect friendship formation.

Although we found significant direct and indirect effects of spatial proximity on friendship, there was little overall effect of spatial proximity on within-school racial friendship segregation. As discussed above, the coefficients on the racial dyad categories decline only slightly when we add data on distance. This suggests that contingent upon mixing together in the same school, within-school-district residential segregation has little effect on friendship patterns. In other words, the strong local effect of distance that we found may explain why a white student is more likely to be friends with another white student who lives within .25 km, but it does not explain his or her tendency to nominate same-race friends who live farther away. Because only 6.2% of friends live within .25 km of one another, the local effect of distance explains little of the overall pattern of social segregation.

Nonetheless, the weak link between within-school residential proximity and racial friendship patterns does not mean that residential segregation is unimportant. On the contrary, residential segregation is important because it results in school segregation, which restricts opportunities for interracial friendship. Recent evidence shows that school segregation closely mirrors patterns and levels of residential segregation in U.S. urban areas (Rickles and Ong 2001). Thus, residential segregation across school districts may be an important source of observed levels of friendship segregation. In this article, we use data from the NCES to calculate the level of middle school and high school segregation in each U.S. metropolitan area. We then combine this information with the estimates of racial friendship homophily in table 6 from the Add Health data to calculate the overall effect of school segregation on friendship segregation. In table 10 we find that residential segregation across schools accounts for about 33%–37% of the overall level of school friendship segregation.

Our results have clear policy implications. In order to reduce school friendship segregation it is important to reduce segregation across schools or school districts. Eliminating residential segregation among the neighborhoods that go to the same school will have only a minor effect. In terms of interracial friendship for middle and high school students, what matters most is whether you have to opportunity to go to the same school.

In sum, our results suggest a classic question of "Is the glass half empty or half full?" On one hand, the results in table 6 show that substantial preferences toward racial friendship segregation exist among students who go to the same school even when we control for spatial proximity; this indicates that residential integration by itself will not make friendship segregation go away. On the other hand, our results indicate that a substantial reduction in friendship segregation is possible if schools were better integrated. Unfortunately, the current trend in school segregation seems to be going the other way. There is ample evidence that school districts across the country are decreasing their efforts to desegregate schools, and that school segregation is rising for blacks and Hispanics even as residential segregation declines overall (Rickles and Ong 2001; Logan 2002; Orfield and Lee 2004; Frankenberg and Lee 2002). Clearly, a debate on the consequences of school segregation would have to include other outcomes such as academic achievement and differences in school quality and funding. Nonetheless, in terms of opportunities for interracial friendship the increasing segregation of public schools in the United States represents a step backward.

APPENDIX A

Modeling Friendship Networks

Modeling friendship networks is complicated because, as shown in the earlier example, friendship choices are influenced by the choices that others make. As a result, the standard statistical assumption of independence does not hold. Friendship ties in a school of g students can be represented by a $g \times g$ matrix, $\mathbf{X}$, where the elements of $\mathbf{X}$ indicate the friendship status of pairs of individuals, that is, $x_{ij} = 1$ if i chooses j as a friend, and 0 otherwise. In exponentiated form, the probability of observing any particular combination of friendships can be defined as

$$\text{prob}\,(X = x^*) = \frac{\exp\,(\boldsymbol{\beta}' z[\mathbf{x}])}{C}, \tag{A1}$$

where $z(\mathbf{x})$ is a vector of variables that affect friendship formation, and $\boldsymbol{\beta}$ is a vector of coefficients (Anderson, Wasserman, and Crouch 1999). C is a constant that ensures that the distribution sums to one, $C = \sum_{\mathbf{x} \in \mathbf{X}} \exp\,[\boldsymbol{\beta}' z(\mathbf{x})]$. The difficulty in estimating equation (A1) is that the normalizing constant C must be calculated over all possible friendship matrices $\mathbf{X}$. Because there are $2^{g(g-1)}$ possible combinations of friends in a school with g students, direct calculation of equation (A1) is impractical if $g \geq 6$ (Strauss and Ikeda 1990). One way to proceed is to simplify by making assumptions about the dependent nature of the data. If we assume that the dyads are independent, then $z(\mathbf{x})$ contains no network variables and equation (A1) turns into a conventional logit model,

$$\left(\frac{\text{prob}\,[x_{ij} = 1]}{\text{prob}\,[x_{ij} = 0]}\right) = \exp\,(\boldsymbol{\beta}' z[\mathbf{x}]). \tag{A2}$$

If we allow the cases to be dependent, but specify the exact nature of the dependency, then we can use equation (A1) to find the odds ratio of the probability of i and j being friends conditional upon the friendship ties in the rest of the data (Wasserman and Pattison 1996; Skvoretz and Faust 1999):

$$\left(\frac{\text{prob}\,[x_{ij} = 1|\mathbf{X}_{ij}^C]}{\text{prob}\,[x_{ij} = 0|\mathbf{X}_{ij}^C]}\right) = \frac{\exp\,(\boldsymbol{\beta}'[\mathbf{z}(\mathbf{x}_{ij}^+)])}{\exp\,(\boldsymbol{\beta}'[\mathbf{z}(\mathbf{x}_{ij}^-)])} = \exp\,(\boldsymbol{\beta}'[\mathbf{z}(\mathbf{x}_{ij}^+) - \mathbf{z}(\mathbf{x}_{ij}^-)]), \tag{A3}$$

where $\mathbf{X}_{ij}^C$ is the friendship matrix for all ties other than x_{ij}, and $\mathbf{z}(x_{ij}^+)$ and $\mathbf{z}(x_{ij}^-)$ indicate the values of the explanatory variables, including network variables, calculated when $x_{ij} = 1$ and $x_{ij} = 0$, respectively.

The conditional log odds in equation (A3) can be estimated by a standard logit model if we assume that we have correctly specified the way in which the dyads are dependent; that is, via mutual friends and/or other more complicated network relationships. In practical terms, calculating the change in the explanatory variables when $x_{ij} = 1$ and $x_{ij} = 0$ in the right hand side of equation (A3) amounts to computing the change in the relevant network variables that depend on the friendship link between i and j. The use of the logit to estimate the conditional log odds in equation (A3) is called the pseudo likelihood "p*" model in the social network literature.

The p* models estimated in this article follow Moody (2001) in using a simple depiction of the role of network structure. The network statistics are defined as follows:

$$\text{the count of the number of friends (choice)} = \sum_{i,j} x_{ij}, \tag{A4}$$

$$\text{the number of reciprocated friendships (reciprocity)} = \sum_{i>j} x_{ij}x_{ji}, \tag{A5}$$

$$\text{the number of 2-out-stars} = \sum_{i,j,k} x_{ij}x_{ik}, \tag{A6}$$

$$\text{the number of 2-in-stars} = \sum_{i,j,k} x_{ij}x_{kj}, \tag{A7}$$

and

$$\text{the number of transitive triads (mutual friends, see below)} =$$

$$\sum x_{ij}x_{jk}x_{ik}. \tag{A8}$$

When the change in the network statistics are calculated for $x_{ij} = 1$ versus $x_{ij} = 0$, the change in the number of 2-out-stars is equivalent to the number of friends i has ("expansiveness," $\sum_k x_{ik}$), the change in the number of 2-in-stars is the number of people who listed k as a friend ("popularity," $\sum_k x_{kj}$), and the change in the number of transitive triads is simply the number of friends that i and j have in common ($\sum_k x_{ik}x_{jk}$). In this simple model, the exogenous independent variables are assumed to have a constant effect on the probability of friendship and are entered into the model by multiplying them by the choice variable. For example, the gender effect $= z(\text{gender}) = \sum_{i,j} x_{ij}(\text{same gender})_{ij}$, where "same gender" is a dummy variable that is 1 if i and j are the same gender. As a result, $z(\text{gender})^+_{ij} - z(\text{gender})^-_{ij} = \text{same gender}_{ij}$. Other models interacting the exogenous independent variables by all the network statistics are possible (Robins, Elliott, and Pattison 2001), but none of the additional interaction terms were significant for the models in this article (results not shown).

Because equation (A3) depicts the conditional probability of friendship given the rest of the network, the logit p^* model of equation (A3) is not a maximum likelihood estimator but a "pseudo likelihood" estimator. The simulation results presented in table 2 help develop an intuition for the performance of the p^* model. In the simulation results in table 2, model 3 estimates p^* models for scenarios 1–4. Model 3 adds an additional network variable, the number of mutual friends connecting individuals i and j. The results in table 2 suggest that, compared to the conventional logit model, the p^* model provides more accurate estimates of the true parameters in the simulated data for all scenarios where there is an endogenous effect because of network structure. In scenario 3, for example, the true parameter for the same-race effect, β_1, is .6, and the p^* estimate in model 3 is .607, compared to an estimate of 1.00 in model 2, which is the conventional logit model assuming the cases are independent.

Although the simulation results in table 2 suggest that the p^* model reduces or eliminates the substantial bias on the estimation of racial homophily present in conventional estimates, because it is a pseudo likelihood model based upon conditional probabilities its estimation properties are not well understood (Snijders 2002; Handcock 2003). Snijders (2002) shows that p^* models tend to underestimate the variance and, in the presence of high correlation among the network variables, may produce inconsistent parameter estimates of the network effects. A proposed alternative to the p^* model is the Markov Chain Monte Carlo (MCMC) method, which uses simulation techniques to estimate the normalizing constant C in equation (1) by simulating $\exp[\boldsymbol{\beta}' z(\mathbf{x})]$ for a sample of possible alternative friendship matrices rather than a complete enumeration. However, the MCMC models are problematic to estimate because of the tendency for the simulated networks **X** to degenerate into full (everyone is a friend) or empty (no one has a friend) matrices (Handcock 2003), and because, under some conditions, it may take a large number of simulations to accurately estimate the coefficients (Snijders 2002). An alternative is a "latent space model," which estimates the network dependency as a latent variable (Hoff, Raftery, and Handcock 2002). At the moment, however, MCMC and latent space models have only been successfully applied to very small data sets, and they are not currently feasible for the type of large data set analyzed in this article. In this article, we adopt a parallel estimation strategy by presenting both p^* and conventional logit results.

APPENDIX B

Alternative Measures of Residential Proximity

Based on the results in table 6, we conclude that within-school district

TABLE B1

Estimated Effects of Alternative Distance Measures on the Likelihood of Friendship

	Model 1: Physical Distance (from Table 6)	Model 2: Log Distance	Model 3: "Relative" Distance	Model 4: Relative Weighted Distance (Random Coefficient)	Model 5: Log Distance (Random Coefficient)
Dyad type:[a]					
White→black	−1.229**	−1.204**	−1.21**	−1.188**	−1.204**
	(.124)	(.126)	(.124)	(.122)	(.124)
Black→white	−1.352**	−1.312**	−1.311**	−1.282**	−1.319**
	(.125)	(.124)	(.122)	(.118)	(.121)
Asian→Hispanic	−.670**	−.652**	−.636**	−.615**	−.622**
	(.186)	(.177)	(.175)	(.177)	(.178)
Hispanic→Asian	−.966**	−.381**	−.966**	−.973**	−.944**
	(.212)	(.139)	(.202)	(.203)	(.205)
Log distance		−.208**			−2.84**
		(.005)			.019
Relative weighted distance			2.401**	.450**	
			(.059)	(.028)	
Suburban school[b]				−.017	.039
				(.031)	(.022)
Rural school[b]				−.086*	.062*
				(.040)	(.029)
Physical distance:[c]					
1–2 km					
.5–1 km	.314**				
	(.027)				
.25–.5 km	.539**				
	(.037)				
Less than .25 km	1.217**				
	(.033)				
All other variables from model 3 of table 6?	Yes	Yes	Yes	Yes	Yes
N	3,857,712	3,857,712	3,857,712	3,857,712	3,857,712

Note.—SEs in parentheses. All models also include all the variables from model 3 of table 6 unless otherwise indicated (coefficients for these variables are not presented here, but are similar to model 3 of table 6).

* $P < .05$.

** $P < .01$.

[a] Excluded category: white→white.

[b] These are level-2 variables predicting the effect of type of school on the coefficient on distance in models 4 and 5. The excluded category is urban school.

[c] Excluded category: >2 km.

residential proximity has a significant effect on friendship probabilities, but that it explains only a small proportion of the overall level of friendship segregation. Table B1 assesses the degree to which our results are robust to alternative measures of distance. The results in table B1 are based on models that include all of the variables in model 3 of table 6, but only selected coefficients for the racial categories are shown.

Model 1 of table B1 repeats the results for the categorical measure of distance used in model 3 of table 6. Model 2 of table B1 uses log distance. Model 3 of table B1 uses relative weighted distance, defined above. In models 4 and 5 of table B1, the effects of relative weighted distance and log distance are estimated as random coefficients, which means that the coefficient is allowed to vary across schools. In models 4 and 5 of table B1, dummy variables for suburban and rural schools are included as explanatory variables for the coefficient on relative weighted distance. In both models, the magnitude of the distance coefficient is smaller in rural schools than in urban schools; this makes sense because rural school districts are more spread out than urban school districts, so that "neighborhood" friends may live farther from each other than in more densely populated urban areas.

The key to interpreting the effect of these different ways of estimating the role of spatial distance on friendship is to compare the coefficients on the racial dyad variables to the corresponding coefficients in model 1 of table B1. The results show that the alternative ways of estimating the effect of distance do not substantially affect the coefficients on white-black or Asian-Hispanic friendship segregation (the results are similar for all the other race categories, available on request). In none of the models can the spatial distance between potential friends who go to the same school be said to be explaining more than about 5% of the effect of race on friendship segregation.

APPENDIX C

Simulating the Predicted Level of Friendship Segregation

The coefficients in the p^* model (model 5 of table 6) estimate underlying homophily preferences, controlling for network structure. As discussed in the text, the overall level of segregation will be higher than the p^* coefficients on race even if race is the only exogenous factor that affects friendship, because of the indirect effect of network propinquity. To estimate the overall level of friendship segregation resulting from the p^* coefficients, we need to simulate the predicted data. To do this, we use a dyad data set of a hypothetical school of 100 students, equally divided between two races. This results in a data set of 9,900 dyads. We use the estimated coefficient on white → black dyads (−.639) and the social network variables transitivity and mutual friends to simulate the predicted data. We estimate 20 simulations, and each simulation is iterated for 10 time periods. The steps involved in each iteration are as follows:

1. Calculate the values of the network variables, reciprocity and mutual friends, from the friendship data of the previous iteration, using equations (6) and (9) in the text.
2. Use the estimated coefficients from model 6 and a random number to calculate the friendship index: Index $= B_1$(same race) $+ B_2$(reciprocity) $+ B_3$(mutual friends) $+ \ln[u/(1-u)]$, where u is a random number between zero and one. The coefficients B_1, B_2, and B_3 are based on the results in model 5 of table 6 and are shown in table B1.
3. Each student is restricted to five friends. For each student, choose the five potential friends with the highest score on the friendship index as "friends."
4. End of iteration. Go to step 1.

An example should make this process clear. If A, B, and C are students in the same school and are all of the same race, then in the first iteration the value of the friendship index for each is .639 + the error term (because there are no friendship ties to start with). If A and C choose B as a friend in the first iteration, and there are no other friendship ties or mutual friends connecting A, B, and C then the value of the friendship index for C choosing A is $.639 + (1) \times (.538) + \ln[u/(1-u)]$ in the second iteration.

Table C1 shows the combined results of the 20 simulations. In model 1, the coefficient on same race is .976. This is the predicted log odds of

TABLE C1
RESULTS OF SIMULATIONS PREDICTING THE LEVEL OF WHITE→WHITE VS. WHITE→BLACK SEGREGATION BASED ON THE RESULTS IN MODEL 5 OF TABLE 6

VARIABLE	COEFFICIENT IN MODEL 6, TABLE 5	MODEL 1		MODEL 2	
		Coefficient	SE	Coefficient	SE
Same race	.639[a]	.976**	.023	.649**	.024
Reciprocity	2.502[b]			2.377**	.026
Mutual friends	.538[b]			.493**	.011
Constant		−3.521**	.019	−4.082**	.022
No. of dyads		198,000		198,000	
No. of simulations		20		20	

[a] Calculated as the log odds of white→white versus white→black friendship using the coefficients in model 5 (0 − [−.639]).
[b] The reciprocity and mutual friends coefficient used in the simulation are calculated from model 6 as $2.79 + 1.84 \times [\ln(100) - 6.16]$ and $.573 + .0268 \times [\ln(100) - 6.16]$, respectively.
** $P < .001$.

TABLE C2
Results of Twenty Simulations of White→Black Segregation

Iteration	Mean	SD	Min	Max
1	.820	.121	.608	1.002
2	.899	.130	.681	1.118
3	.944	.103	.833	1.120
4	.960	.107	.823	1.108
5	.964	.089	.813	1.118
6	.956	.054	.881	1.054
7	.992	.082	.881	1.108
8	1.030	.135	.872	1.287
9	.971	.089	.842	1.184
10 (final) ...	.977	.092	.833	1.151

racial segregation based upon endogenous network structure in a school where all the students are identical except for race and the underlying preference for same-race friends is .639. Table C2 shows the average observed log odds of same-race friendship by iteration for the 20 simulations. This table shows that the observed level of segregation stabilizes after iteration 4, suggesting that the predicted level of segregation is not being underestimated by limiting the simulations to 10 iterations.

APPENDIX D

Decomposing School Friendship Segregation

In this appendix we derive methods for decomposing school friendship segregation into components as a result of school segregation and homophily preferences.

Assuming Constant Homophily Preferences across Schools

If racial preferences for friends are assumed to be constant across schools, we can derive the decomposition of friendship segregation directly. In table D1, p_{ij} is the probability that a randomly selected friend of student k (of race i) would be of race j if the school that k went to was equally divided among students of each race.[20] The log odds ratios correspond to the homophily parameters estimated in table 6. For example, the log odds ratio of a white → black versus a white → white dyad being friends is $\ln(p_{12}/p_{11}) = \alpha_{12} - \alpha_{11}$, which is estimated as -1.318 in model 1 of table 6.

[20] The probability p_{ij} is calculated from the homophily coefficients in table 6 by solving the following equations: $\sum_j p_{ij} = 1$, and $\ln(p_{ij}/p_{ii}) = \alpha_{ij} - \alpha_{ii}$ for all j where α_{ij} is the estimated homophily coefficient for $i \rightarrow j$ dyads.

TABLE D1
HOMOPHILY PREFERENCES

	Probability of Friendship for a Dyad of Potential Friends	
	White	Black
White	p_{11}	p_{12}
Black	p_{21}	p_{22}

TABLE D2
SCHOOL SEGREGATION

	Average Proportion of Students in School, by Race of Student	
	White	Black
White	s_{11}	s_{12}
Black	s_{21}	s_{22}

In addition to homophily preferences, the uneven distribution of students across schools affects interracial friendship by constraining the possibility of social interaction. As discussed in the text, residential segregation of metropolitan areas means that different race and ethnic groups will be distributed unevenly across schools. Table D2 depicts school segregation, measured as the average proportion of each race in the typical white and black student's school, for example, $s_{11} = \sum_j w_j q_{wj} / \sum_j w_i$, where w_j is the number of white students in school j, and q_{wj} is the proportion of the students in school j who are white. The cells in table D2 represent the distribution of dyads of potential friends; for students of race i, the proportion of their dyads that are race $i \rightarrow j$ is s_{ij}. A measure of segregation based on school composition is defined as the odds ratio of the probabilities in table D2:

$$\text{School segregation} = \ln\left(\frac{s_{ij}s_{ji}}{s_{ii}s_{jj}}\right) = S_{ij}. \tag{D1}$$

The predicted probabilities of friendship can be obtained by combining the underlying homophily preferences with school segregation. If p_{ij} is the underlying probability of friendship, and s_{ij} is the proportion of race i's potential friends who are of race j, then the predicted proportion of i's friends who are of race j is

$$\text{Predicted friendship probability} = f_{ij} = \frac{s_{ij}p_{ij}}{\sum_j s_{ij}p_{ij}} \tag{D2}$$

for all racial groups j (i.e., white, black, Hispanic, Asian). As a result, the measure of overall black-white friendship segregation combines the effects of residential segregation across schools and homophily preferences; this is expressed as the odds ratio of interracial friendship:

$$\mathrm{Seg}_{ij} = \ln \frac{\frac{s_{ji}p_{ji}}{\sum_k s_{jk}p_{jk}} \times \frac{s_{ij}p_{ij}}{\sum_k s_{ik}p_{ik}}}{\frac{s_{ii}p_{ii}}{\sum_k s_{ik}p_{ij}} \times \frac{s_{jj}p_{jj}}{\sum_k s_{jk}p_{jk}}} = \ln\left[\left(\frac{s_{ji}s_{ij}}{s_{ii}s_{jj}}\right)\left(\frac{p_{ij}}{p_{ii}}\right)\left(\frac{p_{ji}}{p_{jj}}\right)\right]$$

$$= S_{ij} + \alpha_{ij} - \alpha_{ii} + \alpha_{ji} - \alpha_{jj}. \tag{D3}$$

Equation (D3) shows that we can calculate the relative effect of residential segregation on school friendship segregation by comparing the log-odds ratio of school composition, S, to the estimated homophily parameters in model 1 of table 6.

Allowing Homophily Preferences to Change across Schools

When we include the effect of h, school racial diversity, the probability of friendship (p_{ij}) varies by school and equation (D3) no longer holds. To calculate the predicted level of friendship segregation in the presence of the school-level effect of h, we predict the number of friends of each racial group as $t_{ij}^k = s_{ij}^k p_{ij}^k$, where k represents the kth school in the data, and p_{ij}^k is calculated as in note 12 above with the addition of the school heterogeneity effect.[21] The odds ratio of interracial friendship allowing for school-specific p_{ij}'s is calculated as equation (7) in the text.

[21] For example, the log odds of a white → black versus a white → white dyad being friends in a school with a heterogeneity index of .35 is

$$\alpha_{12} + \delta_{12}(h - .46) - [\alpha_{11} + \delta_{11}(h - .46)] =$$

$$-1.28 - 3.11(.35 - .46) + 0 + .347(.35 - .46) = -.977.$$

Recall that the effect of h is mean standardized.

REFERENCES

Anderson, C. J., S. Wasserman, and B. Crouch. 1999. "A P* Primer: Logit Models for Social Networks." *Social Networks* 21:37–66.

Bearman, Peter S., Jo Jones, and J. R. Udry. 1997. "Connections Count: Adolescent Health and the Design of the National Longitudinal Study of Adolescent Health." Technical report. http://www.cpc.unc.edu/addhealth.

Blau, P. 1977. *Inequality and Heterogeneity*. New York: Free Press.

Charles, C. Z. 2000. "Residential Segregation in Los Angeles." Pp. 167–220 in *Prismatic Metropolis: Inequality in Los Angeles*, edited by L. Bobo, M. L. Oliver, J. H. Johnson, Jr., and A. Valenzuela, Jr. New York: Russell Sage.

———. 2003. "The Dynamics of Racial Residential Segregation." *Annual Review of Sociology* 29:167–207.

Cutler, D., E. Glaeser, and J. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy* 107:455–506.

Damico, Sandra Bowman, and Christopher Sparks. 1986. "Cross-Group Contact Opportunities: Impact on Interpersonal Relationships in Desegregated Middle Schools." *Sociology of Education* 59:113–23.

DuBois, David L., and Barton J. Hirsch. 1990. "School and Neighborhood Friendship Patterns of Blacks and Whites in Early Adolescence." *Child Development* 61:524–36.

Dutton, Susanne E., Jefferson A. Singer, and Ann Sloan Devlin. 1998. "Racial Identity of Children in Integrated, Predominantly White, and Black Schools." *Journal of Social Psychology* 138:41–53.

Ellison C., and D. A. Powers. 1994. "The Contact Hypothesis and Racial Attitudes among Black Americans." *Social Science Quarterly* 75 (2): 385–400.

Falah, Ghazi. 1996. "Living Together Apart: Residential Segregation in Mixed Arab-Jewish Cities in Israel." *Urban Studies* 33:823–57.

Farley R., and W. H. Frey. 1994. "Changes in the Segregation of Whites from Blacks during the 1980s: Small Steps toward a More Integrated Society." *American Sociological Review* 59:23–45.

Festinger, L., S. Schachter, and K. Back. 1950. *Social Pressure in Informal Groups*. New York: Harper.

Fischer, Claude S., et al. 1977. *Networks and Places: Social Relations in the Urban Setting*. New York: Free Press.

Fisher, Bradley J., and David J. Hartmann. 1995. "The Impact of Race on the Social Experience of College Students at a Predominantly White University." *Journal of Black Studies* 26:117–33.

Frank, O., and D. Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81:832–42.

Frankenberg, Erica, and Chungmei Lee. 2002. "Race in American Public Schools: Rapidly Resegregating School Districts." Cambridge, Mass.: Harvard Civil Rights Project. http://www.civilrightsproject.harvard.edu/research/deseg/reseg_schools02.php.

Giordano, Peggy C. 2003. "Relationships in Adolescence." *Annual Review of Sociology* 29:257–81.

Glaeser, E., and J. Vigdor. 2001. *Racial Segregation in the 2000 Census: Promising News*. Washington, D.C.: Brookings Institution.

Grannis, R. 1998. "The Importance of Trivial Streets: Residential Streets and Residential Segregation." *American Journal of Sociology* 103:1530–64.

Hallinan, Maureen T., and Ruy A. Teixeira. 1987. "Students' Interracial Friendships: Individual Characteristics, Structural Effects, and Racial Differences." *American Journal of Education* 95:563–83.

Hallinan, Maureen T., and Richard A. Williams. 1987. "The Stability of Students' Interracial Friendships." *American Sociological Review* 52:653–64.

———. 1989. "Interracial Friendship Choices in Secondary Schools." *American Sociological Review* 54:67–78.

Handcock, Mark S. 2003. "Assessing Degeneracy in Statistical Models of Social Networks." Technical report. University of Washington, Center for Statistics and the Social Sciences. http://www.csss.washington.edu/Papers/wp39.pdf.

Hirschman, Charles, Richard Alba, and Reynolds Farley. 2000. "The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future." *Demography* 37 (3): 381–93.

Hoff, P. D., A. E. Raftery, and M. S. Handcock. 2002. "Latent Space Approaches to Social Network Analysis." *Journal of the American Statistical Association* 97: 1090–98.

Holland, P. W., and S. Leinhardt, S. 1981. "An Exponential Family of Probability Distributions for Directed Graphs (with Discussion)." *Journal of the American Statistical Association* 76 (373): 33–50.

Huckfeldt, R. Robert. 1983. "Social Contexts, Social Networks, and Urban Neighborhoods: Environmental Constraints on Friendship Choice." *American Journal of Sociology* 89:651–69.

Jackman, Mary R., and Marie Crane. 1986. "'Some of My Best Friends Are Black . . .': Interracial Friendship and Whites' Racial Attitudes." *Public Opinion Quarterly* 50:459–86.

Johnson, Monica Kirkpatrick, and Margaret Mooney Marini. 2000. "The Opportunity for Interracial Contact and Interracial Interaction and Friendship." Paper presented at the annual meeting of the American Sociological Association, Washington, D.C., August.

Joyner, K., and G. Kao. 2001. "School Racial Composition and Adolescent Racial Homophily." *Social Science Quarterly* 81 (3): 810–25.

Kandel, D. B. 1978. "Homophily, Selection, and Socialization in Adolescent Friendships." *American Journal of Sociology* 84 (2): 427–36.

Kubitschek, Warren N., and Maureen T. Hallinan. 1998. "Tracking and Students' Friendships." *Social Psychology Quarterly* 61:1–15.

Latane, B., J. H. Liu, A. Nowak, M. Bonevento, and L. Zheng. 1995. "Distance Matters: Physical Space and Social Impact." *Personality and Social Psychology Bulletin* 21: 795–805.

Logan, John. 2002. "Choosing Segregation: Racial Imbalance in American Public Schools, 1990–2000." Report. University at Albany, Lewis Mumford Center for Comparative Urban and Regional Research.

Lubbers, Miranda J. 2003. "Group Composition and Network Structure in School Classes: A Multilevel Application of the p* Model." *Social Networks* 25:309–32.

Marsden, P. V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52 (1): 122–31.

———. 1993. *American Apartheid*. Cambridge, Mass.: Harvard University Press.

Mayhew, Bruce H., J. Miller McPherson, Thomas Rotolo, and Lynn Smith-Lovin. 1995. "Sex and Race Homogeneity in Naturally Occurring Groups." *Social Forces* 74:15–52.

McPherson, M., L. Smith-Lovin, and J. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.

Moody, J. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107:679–716.

Orfield, G., and C. Lee. 2004. *Brown at 50: King's Dream or Plessy's Nightmare?* Cambridge, Mass.: Harvard Civil Rights Project. http://www.civilrightsproject.harvard.edu/research/reseg04/brown50.pdf. Accessed February 29, 2004.

Patchen, M. 1982. *Black-White Contact in Schools: Its Social and Academic Effects*. West Lafayette, Ind.: Purdue University Press.

Pattison, P., and S. Wasserman. 1999. "Logit Models and Logistic Regressions for Social

Networks, II: Multivariate Relations." *British Journal of Mathematical and Statistical Psychology* 52:169–93.

Quillian, Lincoln, and Mary E. Campbell. 2003. "Beyond Black and White: The Present and Future of Multiracial Friendship Segregation." *American Sociological Review* 68:540–66.

Raudenbush, Stephen W., and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis*, 2d ed. Thousand Oaks, Calif.: Sage.

Rickles, Jordan, and Paul M. Ong. 2001. "Relationship between School and Residential Segregation at the Turn of the Century." Working paper. University of California, Los Angeles, School of Public Policy and Social Research. http://lewis.sppsr.ucla.edu/special/metroamerica/factsheets/LCMetroAm_DiscussPaper_2.pdf.

Robins, G., P. Elliott, and P. Pattison. 2001. "Network Models for Social Selection Processes." *Social Networks* 23:1–30.

Schofield, Janet Ward. 1982. *Black and White in School: Trust, Tension, or Tolerance?* New York: Praeger.

Shipler, D. K. 1997. *A Country of Strangers: Blacks and Whites in America*. New York: Random House.

Shrum, W., N. H. Cheek, and S. Hunter. 1988. "Friendship in School: Gender and Racial Homophily." *Sociology of Education* 61:227–39.

Sigelman, L., T. Bledsoe, S. Welch, and M. W. Combs. 1996. "Making Contact? Black-White Social Interaction in an Urban Setting." *American Journal of Sociology* 101 (5): 1306–32.

Sigelman, L., and S. Welch. 1993. "The Contact Hypothesis Revisted: Black-White Interaction and Positive Racial Attitudes." *Social Forces* 71:781–95.

Skvoretz, John, and Katherine Faust. 1999. "Logit Models for Affiliation Networks." *Sociological Methodology* 29:253–80.

Smith, T. W. 1999. "Trendlets: Inter-racial Friendships [and] the Emerging 21st Century American Family." *GSS News* 13:3–4.

Snijders, Tom. 2002. "Markov Chain Monte Carlo Estimation of Exponential Random Graph Models." *Journal of Social Structure* 3, http://www.cmu.edu/joss/content/articles/volindex.html.

Strauss, D., and M. Ikeda. 1990. "Pseudolikelihood Estimation for Social Networks." *Journal of the American Statistical Association* 85:204–12.

Tuma, Nancy Brandon, and Maureen T. Hallinan. 1979. "The Effects of Sex, Race, and Achievement on Schoolchildren's Friendships." *Social Forces* 57:1265–85.

Verbrugge, Lois M. 1983. "A Research Note on Adult Friendship Contact: A Dyadic Perspective." *Social Forces* 62:78–83.

Wasserman, S., and P. Pattison. 1996. "Logit Models and Logistic Regressions for Social Networks, I: An Introduction to Markov Graphs and P*." *Psychometrika* 60:401–26.

Way, Niobe, and Lisa Chen. 2000. "Close and General Friendships among African American, Latino, and Asian American Adolescents from Low-Income Families." *Journal of Adolescent Research* 15:274–301.

White, M. 1987. *American Neighborhoods and Residential Differentiation*. New York: Russell Sage.